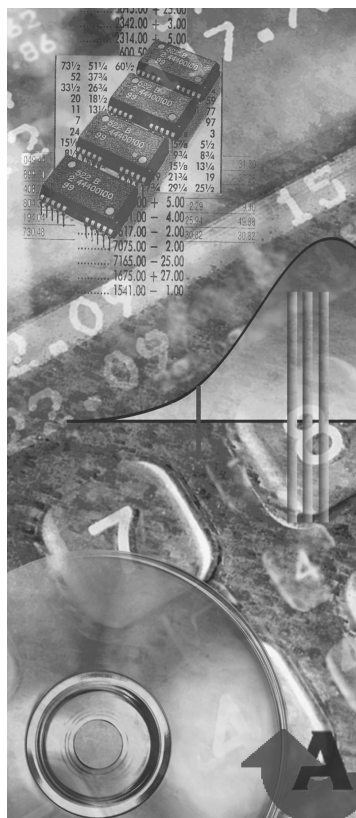


Доверительные интервалы: допущение о неточности оценок

Существуют два пути получения более высокой прибыли: увеличение дохода и уменьшение расходов. В медицинском страховании управлять уровнем доходов сложно, так как страховые компании и правительство устанавливают максимальную сумму, которую они могут заплатить в качестве компенсации при данном диагнозе. Рассмотрим больницу, менеджеры которой пытаются найти разумный ответ на вопрос: “Как много денег в расчете на одного пациента кардиохирургии мы можем заработать или потерять в долгосрочном плане?”.



Тщательный анализ выборки из медицинских и финансовых отчетов о 35 пациентах кардиохирургии показал, что средний доход составляет \$390,26 со стандартным отклонением \$450,56. Итак, об этих 35 пациентах информации достаточно. Но что известно о пациентах кардиохирургии в целом? Вас интересуют не только эти пациенты, а что произойдет в будущем. Вспомним, что стандартная ошибка $450,56 / \sqrt{35} = 76,16$ приблизительно показывает, насколько отличается среднее выборки, равное \$390,26, от среднего генеральной совокупности, из которой эта выборка взята. Однако необходимо двигаться дальше, поскольку такой “аппроксимации” недостаточно. Вам необходимо сформулировать определенное, заслуживающее доверие, точное утверждение. Именно для этой цели используют доверительные интервалы. В данном случае доверительный интервал (способы его вычисления будут описаны ниже) позволяет утверждать следующее.

Мы на 95% уверены, что средний доход в расчете на одного пациента для совокупности, из которой взята выборка из 35 пациентов, находится в пределах от \$235,51 до \$545,01.

Это утверждение о генеральной совокупности определяет область вокруг выборочного значения среднего, \$390,26, является действительно точным утверждением и отражает случайность выборки. Мы не изучали всю генеральную совокупность (которая включает намного большее количество пациентов кардиохирургии), и у нас нет в этом необходимости. Тем не менее мы можем сделать такое же утверждение, как могли бы получить, если бы потратили деньги на анализ и обобщение данных из старых отчетов. Доверительный интервал демонстрирует важную связь между доступным для нас исследованием 35 пациентов и гораздо большей генеральной совокупностью, и эта связь выходит за пределы приближенной интерпретации стандартной ошибки.¹

Практическим следствием такого доверительного интервала является то, что при распространении выводов за пределы данных об исследованных пациентах значение выборочного среднего \$390,26 оказывается не настолько точным, как казалось вначале. Может быть, ошибка в размере \$100 на одного пациента (причем отклонения могут быть в обоих направлениях). Почему ошибка так велика? Причина в изменчивости (одни пациенты приносят прибыль больше, чем другие) и в малом объеме выборки (исследовав данные более чем о 35 пациентах, мы получим более точную оценку среднего генеральной совокупности).

Доверительный интервал можно также использовать, чтобы показать, насколько точно выраженная в процентах доля признака в выборке отражает интересующую нас долю признака в совокупности. Например, результаты маркетингового опроса 150 человек, случайно отобранных из вашей целевой группы, показали, что 46 человек, или 30,7%, знают вашу торговую марку. Вы, конечно же, не верите, что точно 30,7% всей целевой группы знают вашу торговую марку, поскольку вам известно, что случайность процесса построения выборки приводит к ошибке, которая приблизительно равна одной стандартной ошибке. В данном случае стандартная ошибка $S_p = 3,76$ процентных единиц показывает приблизительное различие между выраженными в процентах долями в выборке и в генеральной совокупности. Доверительный интервал, который вычисляют с помощью приведенных далее в этой главе методов, формализует это понятие *приблизительной разности* и позволяет сделать такое заключение.

Мы уверены на 95%, что доля людей, которым известна наша торговая марка, в нашей целевой группе (в генеральной совокупности) находится где-то между 23,3 и 38,0%.

Цель использования доверительных интервалов заключается в том, чтобы по возможности избавиться от неопределенности и сделать как можно более точный вывод. Вероятность дает нам возможность формулировать точные утверждения в условиях неопределенности. Статистика дает возможность извлекать необходимую информацию из данных выборки. Процесс обобщения данных выборки, который приводит к вероятностным утверждениям о всей генеральной совокупности

¹ А что произойдет в том случае, если все записи о пациентах уже занесены в компьютер? Тогда легче будет анализировать более крупные выборки. Однако всегда будет расхождение между пациентами, которых вы наблюдали (выборка), и тем, что будет, вероятно, происходить в будущем. Доверительный интервал будет оставаться полезным и в таком случае, поскольку он фиксирует случайный компонент этой разности. Однако следует также учитывать и систематические изменения на рынке или в технологиях.

сти, называют **статистическим выводом**. В частности, **доверительным интервалом** называют такой вычисленный на данных интервал, который с *известной вероятностью* содержит интересующий нас неизвестный параметр генеральной совокупности, и эта вероятность определяется с учетом случайного эксперимента, который начинается с извлечения случайной выборки. Таким образом, определить доверительный интервал — это лучшее, что можно сделать в условиях неопределенности: это точное вероятностное утверждение вместо неясных замечаний типа: “Мы не уверены, но...” или “Это значение, вероятно, близко к...”.

Доверительные интервалы используют часто, и ниже приведен краткий предварительный обзор их полезных свойств. Есть возможность выбирать вероятность утверждения. Эту вероятность называют **доверительным уровнем** (используют также термины “коэффициент доверия” и “доверительная вероятность”. — *Прим. ред.*). Традиционно его устанавливают равным 95%, но часто используют также значения 90, 99 и даже 99,9%. Платой за более высокий доверительный уровень является более широкий, а значит, и менее полезный интервал. Доверительный интервал для процентного содержания в генеральной совокупности можно легко вычислить, используя стандартную ошибку для биномиального распределения. В зависимости от необходимости можно использовать двусторонний (между двумя значениями) или односторонний (по крайней мере больше, чем некоторое значение) доверительный интервал. Как всегда, следует быть осторожным с не всегда декларируемыми явно, но необходимыми предварительными техническими условиями (в данном случае это нормальность и случайность выборки), поскольку если эти условия не удовлетворяются, то сформулированные на основе доверительных интервалов выводы будут неверными. Необходимо также тщательно различать вероятность 95% для *процесса* построения доверительного интервала и 95% доверительный уровень для конкретного вычисленного интервала.

Сформулируем приблизительное универсальное утверждение о доверительном интервале, которое применяют во многих ситуациях. Если вы с помощью соответствующей несмещенной оценки оценили параметр генеральной совокупности и вычислили соответствующую стандартную ошибку этой оценки, то утверждение о доверительном интервале (в обобщенном виде) можно сформулировать следующим образом.

Приблизительное утверждение о доверительном интервале

Мы на 95% уверены, что параметр генеральной совокупности находится между значением оценки минус две стандартные ошибки и значением оценки плюс две стандартные ошибки.

Следует помнить, что значение нормально распределенной переменной находится в пределах двух стандартных отклонений от своего среднего приблизительно в 95% случаев; вот откуда (отчасти косвенно) возникли эти значения в обобщенной формулировке утверждения о доверительном интервале.

Насколько широко можно применять понятие доверительного интервала? По существу, любое число, которое вы встречаете в газетах, каких-то ваших конфиденциальных стратегических внутренних документах или в телепередачах, является оценкой некоторого важного значения. По сути, все эти оценки имеют соб-

ственные “личные” стандартные ошибки, характеризующие их точность. Знание этих двух величин (оценка и ее стандартная ошибка) позволяет использовать указанное выше приблизительное утверждение о доверительном интервале. Однако далее мы рассмотрим ряд деталей, особенностей и ограничений, которые нужно учитывать в конкретных случаях.

9.1. Доверительный интервал для среднего значения и для доли признака в генеральной совокупности

Мы только что извлекли выборку и с целью оценить среднее генеральной совокупности вычислили выборочное среднее \bar{X} . Пусть нам известно (обычно неизвестно) среднее значение генеральной совокупности, μ , как это показано на рис. 9.1.1.

Расстояние между средним выборки и средним совокупности (ошибка оценивания) *не зависит* от того, будем ли мы использовать в качестве начальной точки измерения этого расстояния выборочное среднее или среднее генеральной совокупности. Иными словами, проведя измерение в единицах стандартной ошибки от среднего совокупности, мы получим тот же результат, который нам даст измерение в единицах стандартной ошибки от выборочного среднего. Это не тривиальная мысль. Поскольку среднее выборки известно, можно проводить измерение в единицах *известного* значения (стандартной ошибки), взяв в качестве начальной точки измерения другое *известное* значение (среднее выборки), и при этом получить

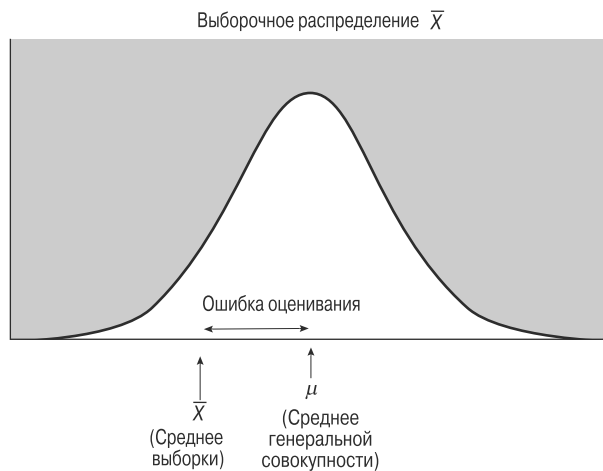


Рис. 9.1.1 Выборочное распределение \bar{X} имеет своим центром значение μ . Расстояние между средним выборки и средним совокупности (ошибка оценивания) одно и то же, независимо от точки отсчета расстояния — от среднего выборки или от среднего совокупности. Доверительный интервал строят, взяв за точку отсчета известное значение \bar{X} , а не неизвестное значение μ

ту же взаимосвязь, как если бы для измерения в качестве начальной точки использовалось (неизвестное) среднее значение генеральной совокупности.

Интуитивные основания построения доверительного интервала следующие. Вспомним, что для нормально распределенной случайной величины вероятность оказаться в пределах двух стандартных отклонений от среднего составляет приблизительно 0,95.² Отсюда получаем следующее вероятностное утверждение.

Вероятность того, что выборочное среднее находится в пределах 1,960 “стандартных отклонений выборочного среднего” от среднего генеральной совокупности, равна 0,95.

Однако это утверждение основано на измерениях относительно *неизвестного* среднего генеральной совокупности, μ . Чтобы избежать этой проблемы, можно провести такие же измерения, но в качестве точки отсчета взять выборочное среднее. В таком случае получим следующее эквивалентное предыдущему вероятностное утверждение.

Вероятность того, что среднее генеральной совокупности находится в пределах 1,960 “стандартных отклонений выборочного среднего” от выборочного среднего, равна 0,95.

Это утверждение все еще включает неизвестный параметр генеральной совокупности, поскольку стандартное отклонение выборочного среднего равно $\sigma_{\bar{x}} = \sigma / \sqrt{n}$. Те, кто занимаются статистикой, часто вместо неизвестных параметров подставляют известные оценки этих параметров. Наше утверждение будет оставаться приблизительно справедливым, если вместо стандартной ошибки использовать значение $S_{\bar{x}} = S / \sqrt{n}$, представляющее наилучшую известную нам информацию о стандартном отклонении среднего выборки. Это приведет к следующему приблизительно вероятностному утверждению:

Вероятность того, что среднее генеральной совокупности находится в пределах 1,960 *стандартных ошибок* от выборочного среднего, *приблизительно* равна 0,95.

К сожалению, это только приблизительно вероятностное утверждение. Чтобы сделать его точным, воспользуемся таблицей *t*-распределения, открытого Стьюдентом (Student) и опубликованного в 1908 г.³ Взяв из *t*-таблицы вместо 1,960 соответствующее критическое значение, получим вместо приблизительно следующего *точное* вероятностное утверждение:

Вероятность того, что среднее генеральной совокупности находится в пределах (критическое значение из *t*-таблицы) стандартных ошибок от выборочного среднего, равна 0,95.

Ценой, которую мы платим за замену неизвестного параметра генеральной совокупности ($\sigma_{\bar{x}}$) соответствующей выборочной оценкой (стандартной ошибкой $S_{\bar{x}}$, вычисленной на имеющихся данных), является то, что критическое значение из *t*-таблицы будет больше, чем 1,960, что, в свою очередь, приведет к более широкому, а значит, менее точному интервалу. Если *n* больше 40, при расчетах вручную можно в качестве приближенного критического значения взять 1,960,

² Можно проверить по таблице нормального распределения, что вероятность находится в пределах 1,960 стандартных отклонений и *точно* равна 0,95.

³ “Стьюдент” — псевдоним У. С. Госсета (W. S. Gossett), управляющего пивоваренной компании Guinness Brewery. Он разработал этот важный метод для контроля за процессом и для улучшения технологии пивоварения.

хотя компьютерные программы, как правило, используют точное, несколько более высокое критическое значение из t -таблицы.

Чтобы получить практически используемый доверительный интервал из этого вероятностного утверждения, следует заменить слова “вероятность 0,95” на слова “доверительный уровень 95%”. Это необходимо сделать в связи с тем, что доверительный интервал на практике формулируют в терминах значений, а не в терминах случайных переменных (более подробно об этом речь пойдет в разделе 9.3). В окончательном виде доверительный интервал (рис. 9.1.2) имеет следующий вид.

Точная формулировка доверительного интервала для среднего генеральной совокупности

Мы на 95% уверены, что среднее генеральной совокупности лежит между значением оценки минус t стандартных ошибок и значением оценки плюс t стандартных ошибок. Таким образом, мы на 95% уверены, что среднее совокупности μ находится где-то между

$$\bar{X} - t S_{\bar{X}} \text{ и } \bar{X} + t S_{\bar{X}},$$

где t взято из t -таблицы для двустороннего доверительного уровня 95%. Шансы на то, что среднее генеральной совокупности находится за пределами этого интервала, равны 5%.

Такой же подход принят при оценивании неизвестной доли (процента) свойства в генеральной совокупности, исходя из доли (процента) свойства в выборке. Это возможно потому, что доля p свойства в выборке представляет собой среднее \bar{X} при условии, что случайная переменная X принимает значения 1 или 0, в зависимости от наличия или отсутствия изучаемого свойства (и, следовательно, процент в генеральной совокупности π также равен среднему совокупности μ). Например, результатом небольшого опроса группы из 5 человек, которым был задан вопрос: “Нравится ли вам цвет этого изделия?”, может быть ряд значений 0, 1, 0, 0, 1, свидетельствующий о том, что цвет изделия нравится второму и пятому из опрошенных. Отсюда, $p = \bar{X} = 0,4$, или 40%. В соответствии с центральной предельной теоремой (которая позволяет использовать нормальное распределение в качестве аппроксимации для биномиального), если n велико, а π не слишком близко к 0 или 1, значения p будут приблизительно нормально распределены и доверительный интервал будет корректным.

Рассмотрим ситуацию, связанную с биномиальным распределением, когда оценивается неизвестная вероятность π некоторого события, на основе наблюдаемых X наступлений этого событий в результате n попыток (при большом

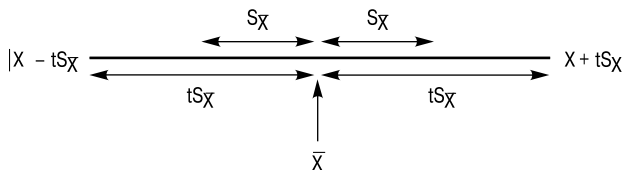


Рис. 9.1.2. 95% доверительный интервал находится в пределах t (приблизительно двух) стандартных ошибок, $S_{\bar{X}}$, в обе стороны от среднего выборки \bar{X}

значении n). Вспомним из главы 7, что для оценки π мы используем долю в выборке, равную $p = X/n$, и что стандартная ошибка p вычисляется по формуле $S_p = \sqrt{p(1-p)/n}$. В соответствии с формулировкой утверждения о доверительном интервале, используя в качестве параметра генеральной совокупности π (вместо μ), в качестве оценки p (вместо \bar{X}) и в качестве стандартной ошибки S_p (вместо $S_{\bar{x}}$), получим следующий доверительный интервал (рис. 9.1.3).⁴

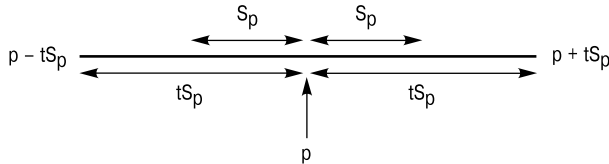


Рис. 9.1.3. Для биномиальной доли π 95% доверительный интервал находится в пределах t (приблизительно двух) стандартных ошибок, S_p , в обе стороны от выборочной доли p

Формулировка доверительного интервала для биномиального распределения (большое значение n)

В частности, мы уверены на 95%, что процент π в генеральной совокупности находится где-то между $p - tS_p$ и $p + tS_p$, где t взято из t -таблицы.

В общем случае длина доверительного интервала определяется, главным образом, размером выборки n и неопределенностью генеральной совокупности. При прочих равных условиях, если размер выборки n большой, то доверительный интервал будет меньше, поскольку формула для вычисления стандартной ошибки включает деление на \sqrt{n} . Уменьшение интервала свидетельствует о меньшей неопределенности из-за наличия большего количества информации. Кроме того, если для выборки неопределенность меньше, то и доверительный интервал будет меньше (это может иметь место при небольшом стандартном отклонении S или, в случае биномиального распределения, если значение доли p близко к 0 или 1).

t-таблица и t-распределение

t -таблицу, представленную здесь под заголовком *таблица 9.1.1*, часто используют для того, чтобы найти множитель при построении доверительного интервала.

⁴ Отметим, что так как в данном случае n велико, значения из t -таблицы будут мало отличаться от соответствующих значений для нормального распределения. Хотя можно использовать непосредственно таблицы для нормального распределения, мы все же используем t по двум причинам: (1) поскольку вместо истинного значения стандартного отклонения σ , используется его оценка S_p ; (2) чтобы сделать подобными процедуры для выборочного среднего \bar{X} и выборочной доли p , поскольку p можно считать содержательно подобным \bar{X} при условии, что отдельные значения закодированы числами 0 и 1. Вычисление стандартной ошибки даст тот же результат, так как при таком кодировании $S_{\bar{x}} = S_p \sqrt{n(n-1)}$.

Таблица 9.1.1. *t*-таблица

Доверительный уровень								
Двусторонний	80%	90%	95%	98%	99%	99,8%	99,9%	
Односторонний	90%	95%	97,5%	99%	99,5%	99,9%	99,95%	
Уровень проверки гипотезы								
Двусторонняя	0,20	0,10	0,05	0,02	0,01	0,002	0,001	
Односторонняя	0,10	0,05	0,025	0,01	0,005	0,001	0,0005	
Для одной выборки: л	В целом: число степе- ней свободы		Критические значения					
2	1	3,078	6,314	12,706	31,821	63,657	318,309	636,619
3	2	1,886	2,920	4,303	6,965	9,925	22,327	31,599
4	3	1,638	2,353	3,182	4,541	5,841	10,215	12,924
5	4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
6	5	1,476	2,015	2,571	3,365	4,032	5,893	6,869
7	6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
8	7	1,415	1,895	2,365	2,998	3,499	4,785	5,408
9	8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
10	9	1,383	1,833	2,262	2,821	3,250	4,297	4,781
11	10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
12	11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
13	12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
14	13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
15	14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
16	15	1,341	1,753	2,131	2,602	2,947	3,733	4,073
17	16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
18	17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
19	18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
20	19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
21	20	1,325	1,725	2,086	2,528	2,845	3,552	3,850
22	21	1,323	1,721	2,080	2,518	2,831	3,527	3,819
23	22	1,321	1,717	2,074	2,508	2,819	3,505	3,792
24	23	1,319	1,714	2,069	2,500	2,807	3,485	3,768
25	24	1,318	1,711	2,064	2,492	2,797	3,467	3,745
26	25	1,316	1,708	2,060	2,485	2,787	3,450	3,725
27	26	1,315	1,706	2,056	2,479	2,779	3,435	3,707

28	27	1,314	1,703	2,052	2,473	2,771	3,421	3,690
29	28	1,313	1,701	2,048	2,467	2,763	3,408	3,674
30	29	1,311	1,699	2,045	2,462	2,756	3,396	3,659
31	30	1,310	1,697	2,042	2,457	2,750	3,385	3,646
32	31	1,309	1,696	2,040	2,453	2,744	3,375	3,633
33	32	1,309	1,694	2,037	2,449	2,738	3,365	3,622
34	33	1,308	1,692	2,035	2,445	2,733	3,356	3,611
35	34	1,307	1,691	2,032	2,441	2,728	3,348	3,601
36	35	1,306	1,690	2,030	2,438	2,724	3,340	3,591
37	36	1,306	1,688	2,028	2,434	2,719	3,333	3,582
38	37	1,305	1,687	2,026	2,431	2,715	3,326	3,574
39	38	1,304	1,686	2,024	2,429	2,712	3,319	3,566
40	39	1,304	1,685	2,023	2,426	2,708	3,313	3,558
Бесконечность		1,282	1,645	1,960	2,326	2,576	3,090	3,291

Сначала обратим внимание на заголовки “Доверительный уровень” и “Двусторонний” в верхней части таблицы. Для построения обычного наиболее распространенного двустороннего 95% доверительного интервала используют соответствующую колонку таблицы. Односторонний интервал будет описан ниже в этой главе, а проверка гипотез — в следующей главе.

В статистике общее понятие **степени свободы** представляет количество независимых элементов информации, используемых для вычисления стандартной ошибки. Для одной выборки число степеней свободы равно $n - 1$ (число на единицу меньше количества наблюдений), так как при вычислении стандартного отклонения из наблюдаемых значений вычитают их среднее⁵. Например, для $n = 10$ наблюдений имеем 9 степеней свободы и поэтому, чтобы получить обычный двусторонний 95% доверительный интервал, используем значение 2,262 из t -таблицы. Если известно точное значение $\sigma_{\bar{x}}$, то используют значение t , равное 1,960, которое соответствует бесконечному числу степеней свободы, поскольку имеется полная информация об изменчивости. Если размер выборки n больше 40, можно в качестве допустимого приближения использовать значение t , соответствующее выборке бесконечного размера (например, 1,960 для 95% уровня доверительности). При этом ваш результат может слегка отличаться от результата компьютерных вычислений, в которых всегда используют точное значение t . Значения t из нижнего ряда таблицы (строка для выборки бесконечного размера) часто называют z -оценкой, так как они соответствуют вероятностям стандартного нормального распределения.

⁵ Вычитание из наблюдаемых значений среднего приводит к потере одной степени свободы, так как сумма полученных таким образом отклонений будет равна 0 и, следовательно, только $n - 1$ отклонений могут свободно изменяться, а последнее отклонение должно быть равно отрицательной сумме остальных.

Каким образом строится t -таблица? Статистики определили t -распределение как выборочное распределение $(\bar{X} - \mu) / S_{\bar{x}}$ при условии, что выборка взята из нормально распределенной совокупности со средним μ . (Это отношение говорит о том, на сколько стандартных ошибок, $S_{\bar{x}}$, среднее выборки \bar{X} превышает среднее генеральной совокупности). Для выборки большого размера (с большим числом степеней свободы) знаменатель близок к $\sigma_{\bar{x}}$, и t -распределение становится близким к стандартному нормальному распределению. Вот почему в нижней части таблицы находятся привычные для нормального распределения значения вероятностей (такие как 1,960). Однако для выборок небольшого размера распределение отличается от нормального (рис. 9.1.4). Знаменатель $S_{\bar{x}}$ влияет на t -распределение таким образом, что кривая распределения имеет более длинную “хвост”, чем при нормальном распределении. Вот почему в верхней части таблицы числа больше, чем в нижней.

Часто используемый 95% доверительный интервал

Почему чаще всего доверительные интервалы вычисляют для уровня доверительности 95%? Говорят, что именно такой выбор рассматривают традиционно как разумный. Уровень 95% представляет собой компромисс между попыткой получить по возможности более высокий уровень доверительности и желанием получить относительно небольшой интервал.

Доверительный интервал для уровня 100%, к сожалению, не очень полезен, так как он слишком велик. Представим себе следующий диалог.

Босс: “Джонс, как Вы думаете, сколько обычный потребитель будет готов платить за нашу новую марку зубной пасты?”

Джонс: “Мы считаем, что обычный потребитель будет готов платить 1,35 доллара за тюбик”.

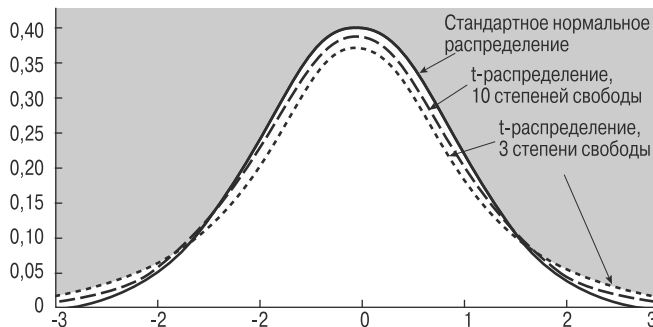


Рис. 9.1.4. t -распределение. Обратите внимание, что по мере роста размера выборки (и, следовательно, числа степеней свободы) форма кривой все больше приближается к стандартному нормальному распределению. Из-за того, что t -распределение характеризуется более длинными “хвостами”, чем стандартное нормальное, приходится продвигаться дальше по кривой, чтобы захватить 95% (или исключить 5%) вероятности. Поэтому при уменьшении числа степеней свободы значения в t -таблице увеличиваются

Босс: “Насколько точна ваша оценка? На что мы можем действительно рассчитывать?”

Джонс: “Отдел анализа на 100% уверен, что типичный потребитель будет готов платить от 0 до 35 миллионов долларов за тюбик”.

Подождите минутку! Это нелепо. Однако дело в том, что для полной уверенности на 100% вам необходимо рассмотреть *все*, даже самые нереальные возможности. Однако продвигаясь в сторону уменьшения от уровня 100% к доверительному уровню, который оставался бы достаточно большим, но оставлял бы некоторый простор для ошибки, можно получить реалистичный и разумный интервал. Давайте попробуем повторить этот диалог еще раз, но несколько иначе.

Босс: “Джонс, как ты думаешь, сколько обычный потребитель будет готов платить за нашу новую марку зубной пасты?”

Джонс: “Мы считаем, что обычный потребитель будет готов платить 1,35 доллара за тюбик”.

Босс: “Насколько точна ваша оценка? На что мы можем действительно рассчитывать?”

Джонс: “Отдел анализа на 95% уверен, что типичный потребитель будет готов платить от 1,26 до 1,44 доллара за тюбик”.

Опыт многих лет свидетельствует, что 95% доверительный уровень является удобным круглым числом, которое достаточно, но в то же время не очень близко к 100%. На практике также используют и другие доверительные уровни: 90, 99 и даже 99,9%; они будут рассмотрены после нескольких примеров.

Пример. Контроль средней толщины бумаги

Контрольно-измерительные приборы, установленные на оборудовании вашей бумажной фабрики, должны быть тщательно отрегулированы для того, чтобы выпускаемая бумага была стандартной толщины. Замеры толщины отобранных листов бумаги (по стандарту толщина бумаги должна равняться 0,004 дюйма) приведены в табл. 9.1.2.

Следует отметить, что средняя величина составила 0,004015 дюйма, что примерно на одну треть процента больше, чем стандартная толщина бумаги, равная 0,004 дюйма. Хотя некоторые отклонения от стандартного значения допускаются практически в любом процессе, вы должны постоянно контролировать состояние оборудования. Невозможно поверить, что среднее значение толщины, равное 0,004015 дюйма, полностью отражает результат работы оборудования. Доверительный интервал позволит вам распространить данные замеров толщины 15 отобранных листов бумаги на всю генеральную совокупность, которую можно считать либо реальной совокупностью (бумага, произведенная за текущий период), либо идеализированной (бумага, которая могла бы быть изготовлена на данном оборудовании в данных условиях). Эта идеализированная совокупность определяет текущее состояние оборудования.

При объеме выборки $n=15$ мы имеем $n-1=14$ степеней свободы и соответствующее значение из t -таблицы для двустороннего 95% доверительного интервала равно 2,145. Доверительный интервал расположен (в предположении о нормальном распределении) между точками

$$\bar{X} - tS_{\bar{x}} = 0,0040146667 - (2,145)(0,0000674986) = 0,00387$$

и

$$\bar{X} + tS_{\bar{x}} = 0,0040146667 + (2,145)(0,0000674986) = 0,00416.$$

Окончательное утверждение о доверительном интервале выглядит так.

“Мы на 95% уверены, что в настоящее время данное оборудование изготавливает бумагу со средней толщиной от 0,00387 до 0,00416 дюйма”.

Полученный доверительный интервал показан на рис. 9.1.5. В качестве результата мы получили точное утверждение с известным уровнем доверительности об общем состоянии оборудования (или же о большом запасе бумаги, из которого была взята выборка), сделанное на основе небольшой выборки.

Таблица 9.1.2. Толщина отобранных листов бумаги (в дюймах)

	0,00385
	0,00358
	0,00372
	0,00418
	0,00380
	0,00399
	0,00424
	0,00375
	0,00449
	0,00422
	0,00407
	0,00434
	0,00381
	0,00421
	0,00397
Среднее	0,0040146667
Стандартное отклонение	0,00026114210
Стандартная ошибка	0,0000674986
n	15

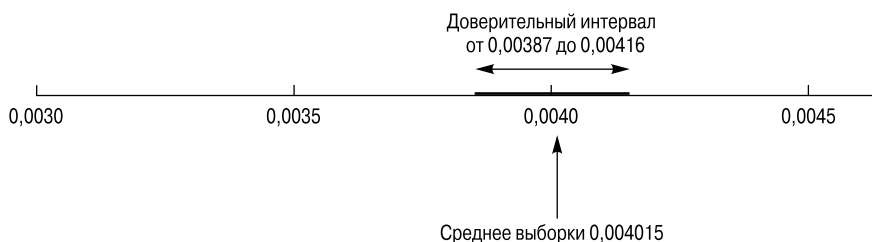
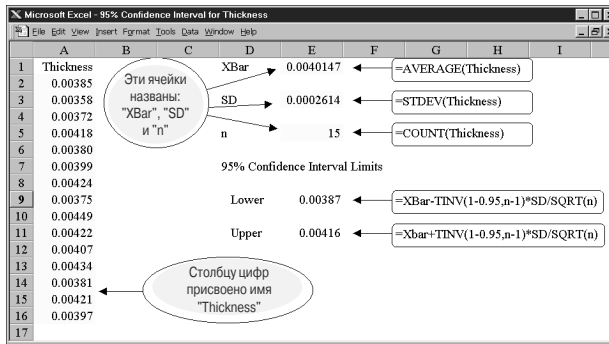


Рис. 9.1.5 Доверительный интервал для средней толщины бумаги, построенный на данных выборки объемом $n = 15$ листов с $\bar{X} = 0,004015$ и $S_{\bar{x}} = 0,0000675$ дюйма

Что можно предпринять, если утверждение недостаточно точно и вы хотите сузить интервал (сделать его меньше, чем от 0,00387 до 0,00416)? Для этого необходимо уменьшить стандартную ошибку $S_{\bar{x}}$ ⁶. Этого можно достичь двумя способами. Во-первых, уменьшить стандартную ошибку можно путем увеличения объема выборки n при прочих неизменных факторах (это легко понять, взглянув на знаменатель формулы $S_{\bar{x}} = S/\sqrt{n}$). Во-вторых, стандартная ошибка уменьшится, если даже не менять объем выборки, но уменьшить изменчивость производственного процесса (из-за числителя формулы $S_{\bar{x}} = S/\sqrt{n}$), определив важные факторы изменчивости и отрегулировав их.

Ниже приведен пример использования электронных таблиц Excel для определения доверительного интервала. Сначала дадим столбцу данных имя Thickness (Толщина), выделив его и воспользовавшись командой Insert⇒Name⇒Define (Вставка⇒Имя⇒Присвоить) из меню Excel. Затем используем функции Excel =AVERAGE(), =STDEV() и =COUNT() (=СРЗНАЧ(), =СТАНДОТКЛОН() и =СЧЁТ()) для вычисления соответственно \bar{X} , S и n и присвоим ячейкам соответствующие имена, чтобы к ним было легко обратиться. 95% доверительный интервал вычислим по формуле $\bar{X} \pm tS/\sqrt{n}$, используя при этом функцию Excel =TINV() (=СТЪЮДРАСПОВР()) для того, чтобы найти t-значение.⁷



Пример. Предвыборный опрос общественного мнения (биномиальный случай)

Президентские выборы 1992 года выглядели как гонка на трех дорожках, когда в отчете опроса *The Wall Street Journal/NBC News* сообщалось, что из 1105 опрошенных зарегистрированных избирателей 33% отдали предпочтение Пэро (независимому кандидату), 31% — Бушу (кандидату от республиканцев) и 28% — Клинтону (кандидату от демократов).⁸ Такие опросы общественного мнения дают ценную информацию для людей, занимающихся организацией предвыборной кампании, а также помогают всем остальным ощущать себя в курсе текущих событий. Однако по различным причинам результаты опросов общественного мнения не являются настолько точными, как это может показаться на первый взгляд. На результаты опроса может оказать влияние то, как сформулирован вопрос (например, в каком порядке перечислены имена в списке,

⁶ Можно попытаться уменьшить каким-то образом t , но это вряд ли поможет, пока исходный объем вашей выборки будет столь небольшой.

⁷ Показанная в данном примере функция =TINV() получает в качестве аргумента “1 — 0,95”, так как эта функция использует именно “единица минус уровень доверительности”, а не сам уровень доверительности. Использовано также выражение “ $n-1$ ”, поскольку функция =TINV() требует в качестве аргумента количество степеней свободы.

⁸ *The Wall Street Journal*, Western Edition, July 9, 1992, p. A1.

или то, задан ли вопрос в положительной или отрицательной форме). Предшествующие вопросы этого же интервью также могут оказывать влияние на результат, формируя некоторый сценарий и вызывая таким образом у опрашиваемого позитивное или негативное мнение. Наконец, существуют статистические ошибки выборки (возможно, как раз их легче контролировать и понять с использованием доверительных интервалов), которые указывают на то, что небольшая выборка не может служить идеальным отражением всей генеральной совокупности.

Наряду с представлением результатов опроса статья также включала следующий взгляд “за кулисы”, взгляд на некоторые детали планирования и анализа результатов таких общенациональных опросов общественного мнения (текст, относящийся к доверительным интервалам, выделен курсивом).

“В ходе опроса *The Wall Street Journal NBC News* две специализирующиеся на проведении опросов фирмы, Peter Hart и Vince Breglio, с воскресенья по среду провели по всей стране телефонные интервью с 1105 зарегистрированными избирателями.

Респонденты опроса отбирались в 263 случайно отобранных географических точках континентальной части США. Каждый регион был представлен пропорционально численности его населения. Домохозяйства отбирались методом, обеспечивающим всем телефонным номерам равные возможности попасть в выборку опроса.

С помощью процедуры, обеспечивающей отбор корректного числа респондентов мужского и женского пола, в каждой семье был опрошен один зарегистрированный избиратель (в возрасте 18 лет или старше). Чтобы данный опрос точно характеризовал всех зарегистрированных избирателей в общенациональном масштабе, результаты исследования были минимально взвешены по профессии респондентов.

Часть вопросов была задана всем респондентам, а часть — только половине. Относительно вопросов, заданных всем, шансы составляют 19 из 20 за то, что результаты данного опроса отличаются от результатов опроса всех имеющих телефон зарегистрированных избирателей США не более чем на три процентные единицы в любом направлении. Предельная ошибка для вопросов, заданных только половине респондентов, составила бы 4,3 процентные единицы. Предельная ошибка для подгруппы зависела бы от размера этой группы”.

В первом абзаце этого текста дан объем выборки $n = 1105$ зарегистрированных избирателей. Второй и третий абзацы поясняют, что в опросе была использована стратифицированная по регионам случайная выборка. Последний абзац описывает доверительные интервалы.

Утверждение “шансы составляют 19 из 20” указывает на вероятность, равную 0,95 или 95% (так как $19/20 = 0,95$). Фраза “результатов опроса всех имеющих телефон зарегистрированных избирателей США” является четким определением генеральной совокупности. Слова “три процентные единицы в каждом направлении” указывают на размер доверительного интервала (tS_p), при этом речь идет о половине длины интервала, которая откладывается в обе стороны от выраженной в процентах оценки доли (p). И наконец, ссылка на предельную ошибку для подгрупп корректно указывает, что подгруппа (например, женщины или юго-западный регион) имеет меньший размер выборки, n , и поэтому здесь может иметь место большая стандартная ошибка и больший доверительный интервал, чем указанные для всей генеральной совокупности три процентные единицы.

Теперь обратите внимание на число 33%, т.е. процент опрошенных избирателей, пожелавших отдать свои голоса за независимого кандидата Пэро. Это число представляет собой точную, выраженную в процентах, долю характеристики в выборке, но в то же время это только выраженная в процентах оценка доли в генеральной совокупности. Двусторонний 95%-ный доверительный интервал находят путем прибавления и вычитания предельной ошибки к выраженной в процентах биномиальной доли. Используя значение $t = 1,960$ из t -таблицы, находим⁹:

$$tS_p = t \sqrt{\frac{p(1-p)}{n}} = 1,960 \sqrt{\frac{0,33(1-0,33)}{1,105}} = 0,0277, \text{ или } 2,77\% .$$

⁹ Здесь предельная ошибка вычисляется так, как если бы речь шла о простой случайной выборке. Для такой стратифицированной выборки, которая использована в нашем примере, необходимо произвести более сложные вычисления.

Как и было заявлено, предельная ошибка, 2,77%, действительно “не превышает три процентные единицы”. 95%-ный доверительный интервал будет находиться в пределах:

от

$$p - tS_p = 0,33 - 0,0277 = 0,302, \text{ или } 30,2\%$$

до

$$p + tS_p = 0,33 + 0,0277 = 0,358, \text{ или } 35,8\%$$

Окончательно утверждение о доверительном интервале можно сформулировать таким образом.

Мы на 95% уверены, что на момент проведения данного опроса от 30,2 до 35,8% имеющих телефон зарегистрированных избирателей США отдали бы свои голоса за Пэро.

Таким образом, исходя из точных данных небольшой по объему выборки мы получили точное утверждение о доверительном интервале для всех зарегистрированных избирателей США. Если быть более осторожным с выводами, то можно уточнить генеральную совокупность, включив в нее тех зарегистрированных избирателей США, до кого в период проведения опроса можно было дозвониться по телефону и кто пожелал выразить свою точку зрения. К тому же, если быть щепетильным, то следует признать, что полученные результаты свидетельствуют о том, как люди определяют свое отношение к Пэро, но не о том, как они будут голосовать.

Другие доверительные уровни

Хотя наиболее часто используемым доверительным уровнем является уровень 95%, иногда используют и другие доверительные уровни. Выбор уровня представляет собой поиск компромисса между размером интервала (меньший интервал является более точным, а значит, и более предпочтительным) и вероятностью того, что интервал включает искомый параметр генеральной совокупности (более высокая вероятность является более предпочтительной). В одних ситуациях необходима очень высокая точность выводов, и тогда увеличивают размеры интервала, чтобы вероятность справедливости утверждения о принадлежности параметра интервалу была выше. В других ситуациях может быть необходим более короткий интервал, и для этого можно допустить, чтобы утверждение о доверительном интервале могло быть неверным более часто. Стандартный 95% доверительный интервал является общепринятым компромиссом между этими двумя факторами, но не единственным решением данной проблемы.

При построении доверительных уровней предпочитают использовать круглые числа (избегая такие сбивающие с толку утверждения, как, например, “быть уверенным на 92,649%”). Значения t-таблицы можно использовать для построения доверительных интервалов для уровней 90, 95 и 99,9% (см. выделенные жирным шрифтом значения в верхней части таблицы), а также для других уровней, которые указаны в таблице прежде всего для построения односторонних интервалов, описанных в разделе 9.4.

Насколько меньше становится доверительный интервал при переходе к более низкому значению доверительного уровня? Для большой выборки относительные размеры доверительных интервалов показаны на рис. 9.1.6.

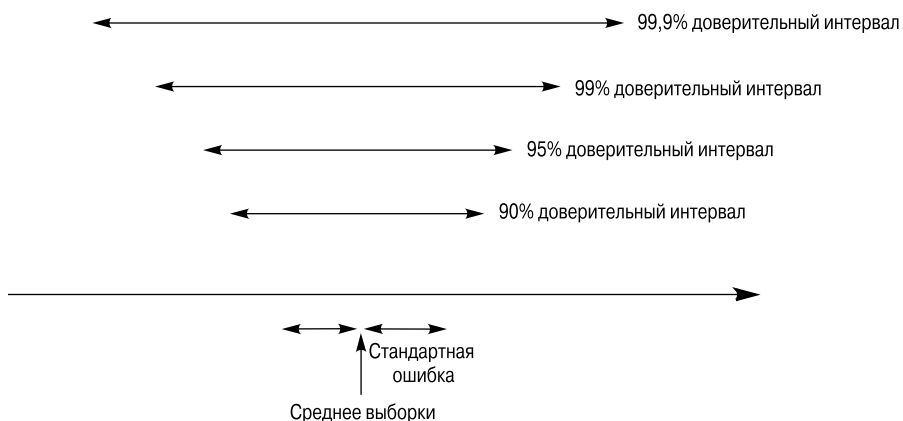


Рис. 9.1.6. Относительные размеры 90, 95 и 99,9% доверительных интервалов для большой выборки. Чем выше требуется уровень доверительности, тем большим должен быть размер интервала, чтобы удовлетворить заявленное требование

Пример. Средняя продажная цена, определенная посредством скидок

Все, наверное, слышали о торговых скидках. Они немного похожи на снижение цены при покупке товара, но для получения скидки необходимо собрать некоторые документы (например, торговый чек; этикетки от бутылок, которые нелегко отодрать, и т.п.), потратить мелкие деньги на отправку письма и, наконец, дожидаться получения чека на один доллар, по которому можно получить наличные в банке.

С точки зрения производителя одно из свойств скидок заключается в том, что они дают определенную полезную информацию. Предположим, что ваша фирма имеет программу скидок на определенный вид батареек, цена которых составляет \$2,39. Вам хотелось бы знать, сколько покупатели действительно заплатили за этот товар после снижения цен на распродажах в магазинах. Тщательно спланированное случайное исследование может дать необходимую информацию, но в данном случае такое исследование может быть не оправдано финансово. Итак, вы решили проанализировать кассовые чеки, присланные вместе с запросами на предоставление скидок.

Прежде всего, какой вид выборки мы имеем в этом случае? Это не случайная выборка в прямом смысле этого слова. Потребители, которые прислали запрос на скидки, не являются репрезентативным "срезом" всех потребителей. В частности, эти потребители могут быть лучше организованы (сохраняют все чеки для того, чтобы потом их выслать) и быть беднее (поэтому для них такая скидка — это деньги, для получения которых стоит побеспокоиться) остальных людей. Тем не менее вы решили изучить генеральную совокупность покупателей, которые высылают запросы на скидки, и рассматривать полученные вами запросы как случайную выборку из этой теоретической генеральной совокупности.

Таким образом, было получено и проанализировано $n = 15\ 603$ кассовых чеков. Итоговые цифры приведены ниже.

Обобщенные данные о кассовых чеках	
n	15 603
Средняя цена продаж	\$2,387
Стандартное отклонение	\$0,318
Стандартная ошибка	\$0,00255

Посмотрите на величину стандартной ошибки! Она так незначительна, потому что размер выборки очень большой. Ваша оценка (\$2,39) очень близка к среднему генеральной совокупности.

95% доверительный интервал, рассчитанный с использованием значения $t = 1,960$, находится между \$2,382 и \$2,392. Таким образом, средняя цена в вашей идеализированной совокупности отличается от полученной оценки \$2,387 не более чем на половину цента.

При такой высокой точности вы немного потеряете, сформулировав утверждение для более высокого доверительного уровня. Например, давайте используем самый высокий доверительный уровень из нашей таблицы. Чтобы достичь доверительного уровня 99,9%, подставим значение $t = 3,291$ вместо 1,960 и получим доверительный интервал

$$\text{от } \bar{X} - tS_{\bar{x}} = 2,387 - (3,291)(0,00255) = \$2,379$$

$$\text{до } \bar{X} + tS_{\bar{x}} = 2,387 + (3,291)(0,00255) = \$2,395.$$

Окончательная формулировка утверждения относительно доверительного интервала будет следующей.

“Мы на 99,9% уверены, что средняя цена покупки батарейки потребителями, которые заинтересованы в получении скидок, находится между \$2,379 и \$2,395”.

Сравним это с 95% доверительным интервалом. Хотя 99,9% доверительный интервал несколько больше, он все еще очень близок к оценке среднего значения цены (приблизительно на цент отличается от \$2,387). Поскольку в этом случае изменчивость (измеренная стандартной ошибкой) очень низка, вы можете сделать очень точное утверждение, которое будет верным с высокой вероятностью.

Пример. Результат производственного процесса

Сейчас, когда введено в действие новое химико-технологическое оборудование, высшее руководство хочет получить надежную информацию о долгосрочных возможностях системы. Технологический процесс очень чувствителен, поэтому, независимо от того, насколько тщательно контролируются параметры, существует некоторое отклонение количества продукции, выпущенной в разные дни или даже в разные часы. Давайте построим доверительный интервал для объема выпускаемой продукции в долгосрочном плане (рассматривая это число как среднее генеральной совокупности), исходя из объемов продукции для выборки временных периодов.

В табл. 9.1.3 приведен столбец данных, состоящий из 12 значений объемов продукции и обычных итоговых показателей. Из таблицы видно, что данные характеризуются сильной изменчивостью.

Поскольку изменчивость данных велика, вы озабочены тем, что доверительный интервал будет больше, чем вам хотелось бы. Вы обсудили эти вопросы с другими сотрудниками и пришли к решению, что 90% доверительный интервал будет приемлемым.

Значение t для двустороннего доверительного интервала при $n - 1 = 11$ степенях свободы равно 1,796. Следовательно, доверительный интервал будет расположен между

$$\bar{X} - tS_{\bar{x}} = 60,3917 - (1,796)(5,4203) = 50,7 \text{ и}$$

$$\bar{X} + tS_{\bar{x}} = 60,3917 + (1,796)(5,4203) = 70,1.$$

Окончательная формулировка утверждения о доверительном интервале будет следующей:

“Мы на 90% уверены, что рассчитанный для большого периода времени средний объем продукции для данного технологического процесса с высокой изменчивостью находится между значениями 50,7 и 70,1 тонны”.

Этот 90% доверительный интервал лишь немного короче 95% доверительного интервала, который имеет границы от 48,5 до 72,3 тонны. Эффект не так уж и велик. Сделав вывод, который может быть неверным в дополнительных 5% случаев, вы получили лишь небольшое увеличение точности в сравнении с общепринятым 95% интервалом.

Таблица 9.1.3. Объем продукции химического оборудования (в тоннах)

	71,7
	46,0
	103,9
	54,4
	43,3
	68,1
	73,4
	45,1
	45,6
	44,9
	77,8
	50,5
Среднее	60,3917
Стандартное отклонение	18,7766
Стандартная ошибка	5,4203
<i>n</i>	12

9.2. Предположения, необходимые для корректного использования

Можно ли быть уверенным в том, что доверительные уровни дают необходимую точность? Когда заявляют о 95% уровне доверительности, можно ли быть уверенным, что с вероятностью 95% среднее генеральной совокупности действительно находится в данном интервале? Для применения статистической теории к конкретному случаю необходимо сделать некоторые технические предположения. Если эти предположения справедливы, доверительные интервалы будут определены правильно. Если предположения не справедливы для конкретной ситуации, то утверждения о доверительных интервалах могут быть неверными.

Что в действительности имеют в виду, когда говорят, что утверждение о доверительном интервале верно? Предположим, что имеется процедура для построения 95% доверительного интервала. Если эта процедура верна и ее повторяют неоднократно (т.е. строят много доверительных интервалов), то приблизительно 95% построенных с ее помощью доверительных интервалов будут содержать среднее генеральной совокупности. Таким образом, у нас нет полной гарантии, что среднее генеральной совокупности определено находится в построенном нами интервале, хотя очень вероятно, что это среднее находится именно там.

Что имеют в виду, когда, в случае некорректности допущения, говорят, что утверждение о доверительном интервале *неверно*? Это означает, что вероятность попадания среднего генеральной совокупности в интервал *не обязательно* равна заявленному уровню доверительности 95% (или другому заявленному уровню

доверительности). Процедура построения должна давать уровень доверительности 95%, но в действительности мы можем получить намного меньший уровень доверительности, равный 50, 10% или еще меньше. Такой доверительный интервал практически бесполезен, даже если он *выглядит* как хороший. С другой стороны, если исходные предположения неверны, реальный уровень доверительности может быть даже *больше* заявленных 95%. К сожалению, в таком случае просто неизвестно, является ли действительный уровень доверительности выше или ниже заявленного уровня 95%.

Для **корректного построения доверительного интервала** необходимо выполнение двух следующих условий: (1) выборка должна быть случайной, (2) распределение должно быть нормальным. Оба этих требования должны быть удовлетворены, чтобы утверждение о доверительном интервале было корректным. Последовательно рассмотрим каждое из этих допущений.

Случайная выборка

Условие 1

Набор данных является случайной выборкой из интересующей нас генеральной совокупности.

Доверительный интервал представляет собой утверждение о среднем генеральной совокупности, основанное на выборочных данных. Естественно, должна быть сильная связь между данными и средним генеральной совокупности. Случайная выборка гарантирует, что ваши данные представляют генеральную совокупность и каждое наблюдение несет в себе новую, независимую информацию. Без случайной выборки нельзя сделать точное вероятностное утверждение о результатах. Если, например, ваша выборка состоит только из ваших друзей, то нельзя ожидать, что рассчитанный вами доверительный интервал будет отражать срез всего общества.

Можно интерпретировать это условие таким образом, что необходимо извлечь случайную выборку из тщательно определенной основы выборки, как об этом говорилось в главе 8. Конечно, в результате таких усилий требуемое условие будет выполнено. Однако это условие не является таким существенным ограничением, как это может показаться. Альтернативный способ удовлетворить необходимое условие о случайности выборки состоит в использовании идеализированной (теоретической) совокупности.

Если имеются некоторые данные и требуется построить доверительный интервал, но эти данные не представляют собой тщательно построенную случайную выборку из четко определенной генеральной совокупности, то можно попытаться построить идеализированную совокупность. Нужно задать себе вопрос о том, что представляют имеющиеся у вас данные. Если вы можете представить себе большую группу и предположить, что ваши данные могут быть рассмотрены как случайная выборка из этой большой группы, тогда обоснованным будет построение доверительного интервала, который даст вам информацию о неизвестном среднем этой идеализированной (теоретической) совокупности.¹⁰

¹⁰ Однако могут быть люди, которые не согласятся с тем, как вы выделили идеализированную (теоретическую) генеральную совокупность. Поскольку это концептуальная, а не чисто статистическая проблема, здесь трудно дать рекомендации.

Например, предположим, что у вас есть данные о людях, недавно обратившихся по вопросу трудоустройства. Строго говоря, эта группа *не* является случайной выборкой из какой-либо совокупности, так как при ее отборе не использовался никакой случайный процесс. У нас недостаточно наблюдений, чтобы утверждать, что эта группа похожа на случайную выборку или что эта группа является достаточно особенной (специфичной). Таким образом, факт остается фактом — в принципе, эта группа не является случайной выборкой. Однако если вы хотите рассматривать этих людей как представителей большей совокупности лиц, ищущих работу на предприятиях, подобных вашему, то вы можете построить доверительный интервал. Этот доверительный интервал будет относиться не к тем конкретным людям, которые обратились по вопросу трудоустройства, а к похожим на них людям из идеализированной (теоретической) совокупности.

Ниже приведен пример неудачного построения доверительного интервала на основе данных, не являющихся случайной выборкой из изучаемой генеральной совокупности.

Пример. Прогноз процентных ставок

Экономические прогнозы позволяют уменьшить неопределенность в процессе стратегического планирования экономической деятельности. Эти прогнозы часто рассматривают как “наилучшую” доступную информацию. Может быть, это и так, хотя неизвестно, насколько в действительности надежны эти данные. Периодически *The Wall Street Journal* публикует старые прогнозы отдельных экономистов вместе с фактическими результатами, чтобы показать, насколько хорошо был сделан прогноз.

На рис. 9.2.1 показана гистограмма прогнозов ставки процента трехмесячных казначейских векселей США в конце 1998 года, которые сделали на 6 месяцев 54 экономиста.¹¹

Двусторонний 95% доверительный интервал, рассчитанный на основе этих 54 прогнозов, покрывает значения ставки процента от 5,01 до 5,15% и не включает фактическое значение, равное 4,46%. Означает ли это, что нам просто не повезло (в том смысле, что в 5% случаев 95% доверительный интервал все же не включает реальное значение), или просто неразумно ожидать, что доверительный интервал может быть использован в такой ситуации? В действительности это не простая неудача; фактически в этой ситуации доверительный интервал интерпретирован неверно, поскольку этот набор данных не является выборкой из той генеральной совокупности, с которой проводится сравнение.

Рассмотрим другой набор прогнозов. На рис. 9.2.2 представлена гистограмма прогнозов ставки процента трехмесячных казначейских векселей США в конце 1984 года, которые сделали на шесть месяцев 22 экономиста.¹²

Двусторонний 95% доверительный интервал, построенный на основе 22 прогнозов, охватывает значения от 10,29 до 10,99% и также не включает фактическое значение ставки процента, равное 7,84%. Фактически прогнозы предсказания очень далеки от действительного значения.

Должен ли нас беспокоить тот факт, что доверительный интервал не включает необходимое нам значение? Необязательно, так как мы знаем, что интервалы включают необходимое нам значение только в 95% случаев. Однако вызывает удивление, что фактическое значение находится чрезвычайно далеко за пределами доверительного интервала. В нашем случае стандартная ошибка равна 0,167% и фактическое значение 7,84% находится на расстоянии $(10,64 - 7,84) / 0,167 = 16,8$ стандартных ошибок от выборочного среднего (10,64%). Это расстояние, равное 16,8 стандартных ошибок, очень велико и требует объяснения.

¹¹ Данные взяты из С. Mitchell Ford, “U.S. Economy Is Seen Growing at Slower Pace in 1999”, *The Wall Street Journal*, January 4, 1999, p.2.

¹² Данные взяты из Т. Herman E. P. Foldessy “Interest Rates to Rise in '85 as Economy Emerges from Doldrums, According to Poll of Economists”, *The Wall Street Journal*, 1985, January 2.

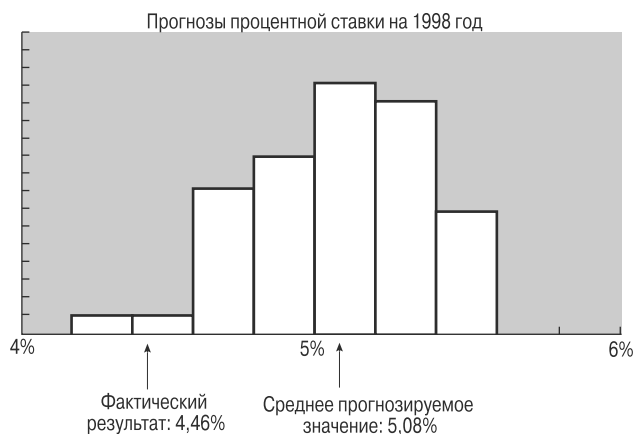


Рис. 9.2.1. Гистограмма прогнозов ставки процента трехмесячных казначейских векселей США в сравнении с фактическими ставками в конце 1998 года

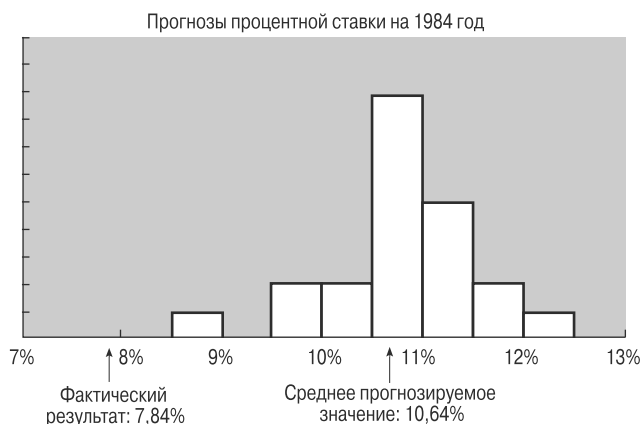


Рис. 9.2.2. Гистограмма прогнозов ставки процента трехмесячных казначейских векселей США в сравнении с фактическими ставками в конце 1984 года. В этом случае прогнозы экономистов не столь согласованы

А объяснение достаточно простое. В данном случае не выполнено условие случайности выборки и, следовательно, нет гарантии, что доверительный интервал является корректным. Таким образом, при планировании будущей деятельности следует более осторожно и критически относиться к прогнозам экономистов.

Однако что же представляют эти экономические прогнозы? Пожалуй, наилучшим решением будет рассматривать эти прогнозы как случайную выборку из идеализированной генеральной совокупности таких же прогнозов ведущих экономистов, сделанных в один и тот же период времени. Тогда такой доверительный интервал будет характеризовать согласованное среднее значение прогнозов этой группы экономистов. Однако этот доверительный интервал не будет характеризовать непосредственно ставку процента, потому что будущее значение ставки процента не является средним генеральной совокупности, из которой взята выборка.

Между прогнозами и фактическим значением ставки процента наблюдается большое различие. Это то, что может произойти и происходит в экономической деятельности. Однако в статистике, если удовлетворены необходимые исходные предположения, корректность соответствующих выводов гарантирована.

Чтобы ход размышлений был лучше понятен, следует отделить предметную область (в данном случае это экономика) от статистических принципов. Лучшее, что можно предпринять в подобной ситуации, — это сделать ограниченный точный статистический вывод и интерпретировать его в соответствии с принятыми в этой предметной области подходами.

Нормальное распределение

Условие 2

Значения данных имеют нормальное распределение.

Теория, лежащая в основе построения доверительного интервала, основана на предположении, что в генеральной совокупности изучаемые данные имеют нормальное распределение. Такое упрощенное условие дает возможность использовать необходимые формулы и построить t -таблицу (что уже сделано для вас). К счастью, есть две причины, по которым на практике это требование не является столь критичным.

Во-первых, никогда нельзя точно сказать, является совокупность полностью нормально распределенной или нет, поскольку имеется только случайная выборка. Поэтому на практике обращаются к гистограмме данных, чтобы удостовериться, что распределение является *приблизительно* нормальным, т.е. не слишком асимметричным и без особо отличающихся экстремальных значений.

Во-вторых, часто помогает центральная предельная теорема. Поскольку статистический вывод основан главным образом на выборочном среднем \bar{X} , необходимо, чтобы в первую очередь выборочное распределение \bar{X} было распределено приблизительно нормально. Центральная предельная теорема утверждает, что если n достаточно велико, \bar{X} будет распределено приблизительно нормально, даже если отдельные элементы в генеральной совокупности (и в выборке) не распределены нормально.

Итак, можно сформулировать следующее правило для практического применения.

Условие 2 (на практике)

Изучите гистограмму данных. Если она выглядит приблизительно нормально, то все в порядке (т.е. доверительный интервал приблизительно корректен). Если гистограмма немного асимметрична, то необходимо иметь не очень малую выборку. Если гистограмма умеренно асимметрична или имеет несколько умеренных отличающихся значений, то необходимо иметь выборку большого размера. Если гистограмма сильно асимметрична или имеет большие экстремальные значения, то это должно вызывать беспокойство.

Для биномиального распределения следствие центральной предельной теоремы заключается в том, что выраженная в процентах доля признака в выборке p при большом значении n распределена приблизительно нормально (при условии, что выраженная в процентах доля признака в генеральной совокупности не

слишком близка к 0% или 100%). Это показывает, каким образом условие нормальности распределения может быть выполнено (приблизительно) для биномиального распределения.

Как быть, если условие нормальности распределения не выполняется совсем, скажем при сильной асимметрии? Одним из способов заключается в поиске и использовании такого преобразования данных (например, логарифмического), которое привело бы к нормальному распределению; однако в таких случаях необходимо помнить, что в результате будет получен доверительный интервал для среднего *логарифмов* значений генеральной совокупности. Другой способ заключается в использовании непараметрических методов, которые будут описаны в главе 16.

9.3. Интерпретация доверительного интервала

Что вы имеете в виду, когда говорите, что исходя из значений веса в выборке из дневной продукции вы на 95% уверены, что средний вес всех изготовленных сегодня упаковок мыла лежит в пределах от 15,93 до 16,28 унций? Это похоже на вероятностное утверждение, но с ним необходимо тщательно разобраться. Средний вес всех выпущенных сегодня упаковок мыла является некоторым конкретным неизвестным числом. Это число либо принадлежит интервалу, либо не принадлежит. А раз так, то откуда появляется вероятность?

Какое событие имеет вероятность 95%?

Чтобы возникла вероятность, должен иметь место случайный эксперимент. Вероятность скорее относится к *процессу* в целом, чем к конкретному результату. Когда вы говорите, что на 95% уверены в том, что среднее значение веса в генеральной совокупности находится в пределах от 15,93 до 16,28 унций, то делаете вывод о точных числовых результатах, исходя из имеющихся данных. Однако вероятность 95% возникает из самого процесса, который рассматривает значения как *случайные*. Более тщательная формулировка вероятностного утверждения может быть такой: “Вероятность события “средний вес в генеральной совокупности находится в пределах доверительного интервала” для случайного эксперимента “случайно отобрать несколько упаковок и построить доверительный интервал” равна 95%”. Каждый раз, когда собирают данные и вычисляют 95% доверительный интервал, проводят случайный эксперимент, который определяет вероятность для каждого события. Вероятность того, что неизвестное среднее генеральной совокупности попадает в вычисленный интервал, равна 0,95.

Тонкостью является вопрос о том, к какому времени относится информация. Можно обоснованно заявить, что вероятность того, что завтра индекс фондового рынка поднимется, равна 55%. Однако, когда завтра после обеда окажется, что индекс действительно пошел вверх, уже не остается ни неопределенности, ни вероятности — индекс пошел вверх. Но перед тем как это произошло, неопределенность *была*. Единственное различие между этим примером о фондовом рынке и обычным построением доверительного интервала состоит в том, что в последнем случае среднее генеральной совокупности либо находится в доверительном интервале, либо нет, однако при этом вы можете никогда не узнать, действительно ли среднее находится в этом интервале или нет!

Чтобы понять, что означает вероятность 95%, полезно представить себе многократное повторение процесса извлечения выборки и построение множества доверительных интервалов, вычисленных на данных различных случайных выборок. Из понятия относительной частоты и закона больших чисел (из главы 6) мы знаем, что приблизительно 95% этих случайных известных интервалов включают конкретное неизвестное среднее значение генеральной совокупности. Это показано на рис. 9.3.1. Обратите внимание, что каждая выборка имеет свое среднее \bar{X} , поэтому некоторые интервалы смещены вправо или влево относительно друг друга. Кроме того, каждый интервал имеет свою стандартную ошибку $S_{\bar{x}}$, поэтому одни интервалы больше, другие — меньше. Заметьте, что даже те доверительные интервалы, которые не содержат среднее генеральной совокупности, тем не менее, расположены к нему довольно близко.

Ваши жизненные достижения

Конечно, обычно в конкретной ситуации вы вычисляете только *один* доверительный интервал для среднего генеральной совокупности. Однако если проводится много подобных исследований, независимых друг от друга (т.е., выборки извлекаются независимо), то можно интерпретировать смысл “95% доверительности” в терминах достижений вашей жизни. Если вы за всю свою жизнь построили много 95% доверительных интервалов и если в каждом случае были удовлетворены необходимые условия, то приблизительно 95% этих доверительных интервалов действительно содержали соответствующие им средние значения генеральных совокупностей.

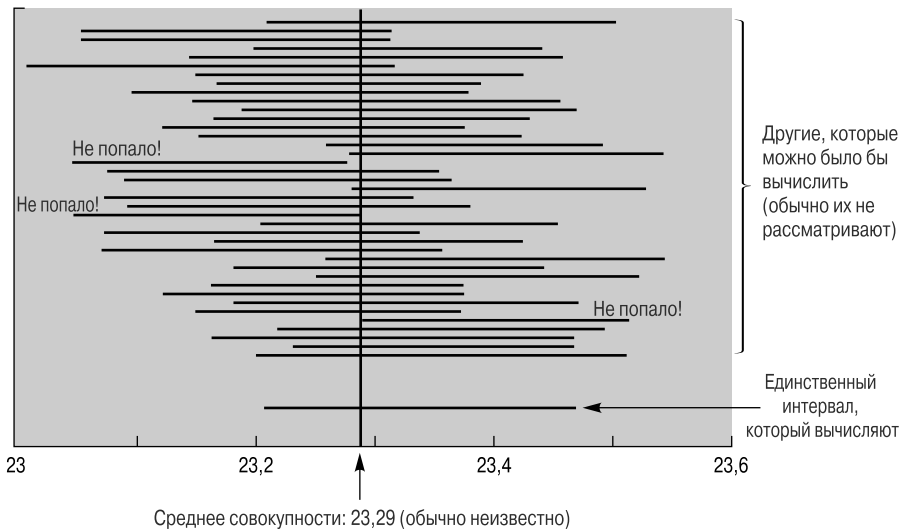


Рис. 9.3.1. Что если использовать другую случайную выборку? Этот рисунок показывает, как могут различаться доверительные интервалы, построенные на основе различных случайных выборок (построенных независимо) из одной и той же генеральной совокупности. Если процесс будет продолжаться достаточно долго, то 95% этих интервалов будет включать неизвестное среднее при условии выполнения исходных предположений

Обучаясь на курсах гольфа в доме престарелых и вспоминая свою жизнь, вы чувствуете удовлетворение от того, что в 95% случаев ваши доверительные интервалы были правильными. К этому чувству примешивается также огорчение, поскольку 5% интервалов были неверны. И вы *никогда не сможете узнать*, в каких случаях результаты вашего труда были верны, а в каких нет! Таковы методы статистического вывода.

9.4. Односторонние доверительные интервалы

В некоторых случаях нет необходимости утверждать, что среднее генеральной совокупности находится *между* двумя границами доверительного интервала. Достаточным может быть утверждение, что среднее генеральной совокупности по крайней мере *не меньше, чем* некоторое число, или (в других случаях) что среднее генеральной совокупности *не больше, чем* некоторое число. **Односторонний доверительный интервал** устанавливает с известной доверительностью, что среднее генеральной совокупности либо *не меньше*, либо *не больше* некоторого вычисленного значения. Если подходить тщательно, то построение одностороннего доверительного интервала может быть даже более эффективным, чем использование двустороннего.

Например, вас может только лишь интересовать, чтобы нечто было *достаточно большим*: “Мы на 95% уверены, что сумма продаж составит, по меньшей мере, \$560 000”. Или вас может только интересовать, чтобы нечто было *достаточно небольшим*: “Мы на 95% уверены, что уровень брака составляет не более, чем один дефект на 10 000 единиц производимой продукции”. В ситуациях такого типа полезны односторонние доверительные интервалы.

Внимание! Не всегда можно использовать односторонний доверительный интервал

Для использования одностороннего доверительного интервала необходимо выполнение следующего важного критерия.

Критерий для использования одностороннего доверительного интервала

Чтобы использовать односторонний доверительный интервал, вы должны быть уверены, что *независимо* от характера данных вы будете использовать односторонний интервал с той же стороны (т.е. “не меньше чем” или “не больше чем”). Если изменение характера данных приводит к необходимости использовать односторонний интервал с *другой* стороны, то вы должны использовать вместо одностороннего двусторонний интервал. Если есть сомнения, используйте двусторонний интервал.

Предположим, что себестоимость единицы производимой вами продукции составляет \$18, у вас есть необходимые выборочные данные и вы готовы вычислять доверительный интервал. Существует соблазн проделать это следующим образом: если оценка издержек *высокая* (больше \$18), вы будете утверждать, что на 95% уверены, что издержки *по меньшей мере...*, а если оценка издержек *невысокая* (меньше \$18), вы заявите, что на 95% уверены, что издержки *не более, чем...* *Не поддавайтесь соблазну!* Потому что в соответствии с указанным выше

критерием изменение стороны доверительного интервала с одной на другую исходя из данных выборки недопустимо. Вам следует строить двусторонний интервал (вы на 95% уверены, что издержки находятся между... и ...). Тому есть две причины. Во-первых, вас интересуют обе стороны доверительного интервала: иногда одна, иногда другая. Во-вторых, при работе с односторонним доверительным интервалом переход от одной стороны к другой может сделать недействительным ваше вероятностное утверждение и таким образом истинный уровень доверительности вашего утверждения может оказаться более низким, чем заявленный 95%.¹³

Вычисление одностороннего доверительного интервала

Чтобы вычислить односторонний интервал, сначала по табл. 9.1.1 найдите значение t , используя графу “Односторонний” для уровня доверительности в верхней части таблицы. (Используйте ту же строку таблицы, что и для двустороннего интервала, поскольку число степеней свободы не изменяется и равно $n - 1$.) Например, чтобы вычислить 95% односторонний доверительный интервал для выборки размером $n = 23$, следует использовать $t = 1,717$. Для 99,9% одностороннего доверительного интервала при $n = 35$ используйте $t = 3,348$.

Далее следует выбрать *одно* из следующих утверждений для одностороннего доверительного интервала.

“Мы на 95% уверены, что среднее генеральной совокупности *не меньше чем* $\bar{X} - t_{\text{односторонний}} S_{\bar{X}}$ ”

или

“Мы на 95% уверены, что среднее генеральной совокупности *не больше чем* $\bar{X} + t_{\text{односторонний}} S_{\bar{X}}$.”

Чтобы легче вспомнить, следует ли вычитать или прибавлять второй член, необходимо убедиться в том, что среднее значение выборки \bar{X} включено в односторонний интервал. (Так должно быть, поскольку это ваша наилучшая оценка среднего генеральной совокупности.) Таким образом, если односторонний доверительный интервал направлен в сторону больших значений (“не меньше чем”), то он должен начинаться *ниже* среднего значения выборки, а если односторонний доверительный интервал направлен в сторону меньших значений (“не больше чем”), то он должен начинаться *выше* среднего значения выборки.

Рис. 9.4.1 иллюстрирует этот факт, а также сравнение одно- и двусторонних доверительных интервалов.

Начальная точка *одностороннего* 95% доверительного интервала совпадает с одной из двух граничных точек *двустороннего* 90% доверительного интервала. Дело в том, что есть две причины, по которым двусторонний интервал может оказаться неверным: среднее совокупности либо слишком велико, либо слишком

¹³ В наихудшем случае при таких изменениях вы можете, заявив уровень доверительности 95%, реально получить 90% односторонний доверительный интервал. Так будет, если вы будете изменять стороны в зависимости от того, больше или меньше значение \bar{X} значения μ . В таком случае у вас будет ошибка 5% в *каждом из* интервалов, и вы получите в качестве общей ошибки не 5%, на которые рассчитываете, а 10%.

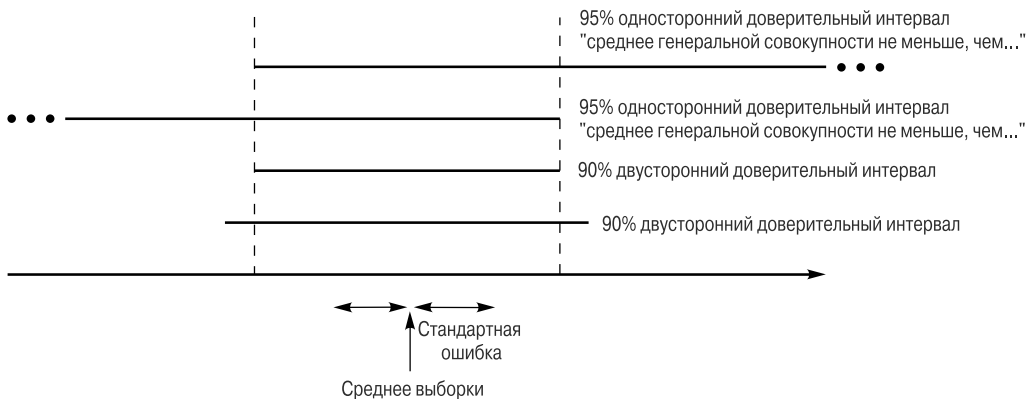


Рис. 9.4.1. В верхней части рисунка показаны оба типа односторонних доверительных интервалов. Односторонний доверительный интервал всегда включает среднее значение выборки, начинаясь с точки по одну сторону от выборочного среднего и продолжаясь до бесконечности в другую сторону. Обратите внимание, что начальная точка 95% одностороннего доверительного интервала совпадает с одной из граничных точек 90% двустороннего доверительного интервала

мало. Односторонний интервал, который имеет общую граничную точку с двусторонним интервалом, может быть неверным только в половине соответствующих случаев.

Односторонний доверительный интервал позволяет сконцентрировать внимание на наиболее интересных случаях. Если вас интересуют ошибки только по одну сторону и совсем не заботят ошибки с другой стороны, тогда односторонний интервал может иметь начальную граничную точку *ближе к среднему значению выборки* (и поэтому он будет точнее), чем двусторонний доверительный интервал. Например, для большой выборки со средним, равным 19,0, и стандартной ошибкой, равной 8,26, вместо того чтобы утверждать, что среднее генеральной совокупности лежит между 2,81 и 35,2, вы сможете сказать, что это среднее не меньше 5,41. Знание, что среднее генеральной совокупности не меньше 5,41, дает больше информации, чем знание, что это среднее не меньше 2,81. Вы можете формулировать более строгое утверждение о нижней границе благодаря тому, что вы не формулируете вообще никаких утверждений о верхней границе.

Пример. Экономия от применения новой системы

Вы оцениваете новую автоматизированную систему производства и намерены приобрести ее, если выяснится, что она обеспечивает достаточную экономию денежных средств на единицу выпускаемой продукции. Вы организовали установку системы с условием, что сможете испытать ее в течение недели. Система должна быть запрограммирована для всего ассортимента типовых изделий, и экономия будет определяться для каждой единицы выпускаемой продукции.

Что в этом случае представляет собой генеральная совокупность? Это теоретическая совокупность всех тех изделий, которые могла бы выпускать система в условиях, аналогичных тем, при которых проходят испытания. Статистический вывод может помочь вам в этом случае, позволяя на основе информации о группе конкретных выпущенных изделий сделать вывод о среднем для значительно большего количества изделий, которые могли бы быть изготовлены в будущем в аналогичных условиях.

Следует ли в этом случае использовать односторонний интервал? Да, следует, поскольку независимо от поведения данных, вас интересует только одно — сможете ли вы сэкономить деньги. Интересующее вас вероятностное утверждение может выглядеть так: “Мы на 95% уверены, что среднее значение сэкономленных средств на единицу продукции при длительном периоде эксплуатации будет не меньше чем...”.

Для выборки размером $n=18$ единиц произведенной продукции со средним значением сэкономленных средств $\bar{X} = \$39,21$ и стандартной ошибкой $S_{\bar{x}} = \$6,40$ односторонний доверительный интервал будет бесконечно простирается в направлении больших значений, начиная с

$$\bar{X} - t_{\text{односторонний}} S_{\bar{x}} = 39,21 - (1,740)(6,40) = 28,07 .$$

Поэтому окончательная формулировка утверждения относительно одностороннего доверительного интервала будет следующей.

Мы на 95% уверены, что среднее значение сэкономленных средств не меньше чем \$28,07 на единицу выпускаемой продукции.

Обратите внимание, что односторонний доверительный интервал включает среднее выборки $\bar{X} = \$39,21$, как и должно быть. Иными словами, выборочное среднее $\bar{X} = \$39,21$ удовлетворяет формулировке утверждения о доверительном интервале, будучи не меньше \$28,07. Было бы неверно использовать другую граничную точку. Такой вид проверки правильности выбора начальной точки обеспечивает корректность построенных односторонних доверительных интервалов.

Чтобы вычислить односторонний доверительный интервал при другом уровне доверительности, просто возьмите в таблице и подставьте в формулу соответствующее t -значение. Например, при построении 99% одностороннего доверительного интервала используют значение $t = 2,567$. По сравнению с 95% интервалом это утверждение относительно значения сэкономленных средств более слабое, но этому утверждению вы доверяете больше.

Мы на 99% уверены, что среднее значение сэкономленных средств составляет не менее \$22,78 на единицу выпускаемой продукции.

Пример. Затраты на командировки

Пытаясь спланировать реалистичный бюджет командировок, вы проанализировали стоимость типичных поездок за последнее время. Чтобы быть уверенным, что бюджет покрывает требования следующего года, вам нужно получить максимальную величину средних затрат (в долларах) на одну командировку. Тогда вы сможете утверждать, что средняя стоимость командировки не будет превышать это значение. Ввиду того, что вас интересует ограничение стоимости командировки только с одной стороны, вы можете использовать односторонний доверительный интервал. Вы выбираете уровень доверительности 95%.

Анализируя список последних 83 командировок, вы определили, что средняя стоимость командировки составляла \$1 286 при стандартной ошибке, равной \$71,03. Односторонний доверительный интервал будет включать все значения, начиная с \$0 (так как величина стоимости не может быть отрицательным числом) и до верхней границы:

$$\bar{X} + t_{\text{односторонний}} S_{\bar{x}} = 1286 + (1,645)(71,03) = \$1403 .$$

Окончательная формулировка утверждения об одностороннем доверительном интервале будет выглядеть таким образом.

Мы на 95% уверены, что среднее значение расходов на одну командировку будет не более \$1403.

Чтобы удостовериться, что в использованной формуле должен стоять именно знак “плюс” (а не “минус”), посмотрим на среднее значение выборки и увидим, что это значение (\$1286) действительно находится в пределах доверительного интервала (\$1286 не превышает \$1403).

Построенный доверительный интервал имеет ограниченное применение, поскольку использованные данные не представляют собой действительно случайную выборку из интересующей вас генеральной совокупности. Вы хотите спрогнозировать будущие командировочные расходы, в то время как данные вашей выборки относятся к прошлому. Доверительный интервал учитывает изменчивость командировочных затрат в прошлом, что является для вас полезной информацией. Однако он не учитывает (и не может учитывать) изменение командировочных расходов в будущем.

9.5. Интервалы предсказания

Доверительный интервал дает информацию о том, где с известной вероятностью находится *среднее генеральной совокупности*. Это очень полезно, если вы ищете обобщающую характеристику для большой генеральной совокупности. Однако если вы хотите получить информацию о наблюдаемом значении для *отдельного случая*, этот доверительный интервал вам не подходит. Вместо этого вам необходим более широкий интервал, отражающий не только оценку неопределенности \bar{X} , равную $S_x = S/\sqrt{n}$ (которая может быть очень маленькой при большом значении n), но и оценку неопределенности S отдельного наблюдения.

Интервал предсказания позволяет использовать данные из выборки для предсказания с известной вероятностью нового наблюдения при условии, что это новое наблюдение получают тем же способом, что и прошлые данные. Таким образом, ситуация заключается в следующем. Имеется случайная выборка размером n из генеральной совокупности, и в результате проведенных для этой выборки измерений получены значения X_1, \dots, X_n . Вам необходимо сделать предсказание относительно *нового* элемента, случайно выбранного из этой же генеральной совокупности.

Используемая здесь мера неопределенности представляет собой **стандартную ошибку предсказания**, т.е. величину изменчивости расстояния между средним значением выборки и новым наблюдением. Здесь объединены два вида случайности: для среднего значения выборки и для нового наблюдения. Стандартную ошибку предсказания находят умножением стандартного отклонения на корень квадратный из выражения $(1 + 1/n)$.

Стандартная ошибка предсказания

$$S\sqrt{1 + \frac{1}{n}}$$

Значение стандартной ошибки предсказания больше оценки S изменчивости отдельных элементов в генеральной совокупности. Так и должно быть, поскольку интервал предсказания должен объединять изменчивость отдельных элементов в совокупности (которая измеряется S) и изменчивость выборочного среднего значения (которая измеряется $S_x = S/\sqrt{n}$).

Вычислив оценку (\bar{X}) и стандартную ошибку предсказания, можно построить интервал предсказания таким же образом, как и обычный доверительный интервал. Также по таблице находят t -значение для заданного уровня доверительности прогноза и размера выборки n (конечно, не включая в выборку до-

полнительное наблюдение). Отличается только формула для вычисления стандартной ошибки; удостоверьтесь, что вы используете формулу для стандартной ошибки предсказания, а не стандартную ошибку среднего значения выборки.

Интервал предсказания для нового наблюдения	
Двусторонний:	<p>“Мы на 95% уверены, что новое наблюдение будет находиться между</p> $\bar{X} - tS\sqrt{1+1/n} \text{ и } \bar{X} + tS\sqrt{1+1/n} \text{ ”.}$
Односторонний:	<p>“Мы на 95% уверены, что новое наблюдение не меньше, чем</p> $\bar{X} - t_{\text{односторонний}}S\sqrt{1+1/n} \text{ ”}$
или	<p>“Мы на 95% уверены, что новое наблюдение не больше, чем</p> $\bar{X} + t_{\text{односторонний}}S\sqrt{1+1/n} \text{ ”.}$

Какой смысл имеет здесь значение 95%? Это вероятность, соответствующая следующему случайному эксперименту: берем случайную выборку, находим интервал предсказания, берем новое случайное наблюдение и смотрим, попадает ли это новое наблюдение в построенный интервал. Обратите внимание, что вероятность 95% относится как к извлечению новой выборки, так и к извлечению нового наблюдения. Это естественно, потому что, поскольку одна выборка отличается от другой, доля новых наблюдений, попадающих в интервал предсказания, будет также разной для разных выборок. Усреднив случайность первоначальной выборки, получим вероятность, равную 95% (или другую необходимую вероятность).

В приведенной ниже таблице показано, когда необходимо использовать интервал предсказания вместо доверительного интервала.

Когда необходимо получить информацию о...	Используют ...
среднем генеральной совокупности	доверительный интервал
новом (таком же, как и остальные) наблюдении	интервал предсказания

Пример. Сколько времени необходимо для выполнения заказа

Когда нужно заказывать комплектующие для производства? Если вы закажете намного раньше, вам придется заплатить процент за кредит, взятый для приобретения комплектующих, и, кроме того, придется платить арендную плату за складское помещение для их хранения. Если вы закажете слишком поздно, вы рискуете остаться без необходимых комплектующих и придется на время остановить производственную линию.

При выполнении последних восьми поставок ваш поставщик говорил: “Следующий заказ будет выполнен через две недели”. Вы отметили для себя, сколько рабочих дней в действительности потребовалось для выполнения последующих заказов, и у вас получились такие цифры:

10, 9, 7, 10, 3, 9, 12, 5

Среднее значение равно $\bar{X} = 8,125$ дня со стандартным отклонением $S = 2,94897$. Стандартная ошибка среднего равна $S_x = 1,04262$, однако не она нас интересует. Стандартная ошибка прогноза вычисляется следующим образом.

$$\begin{aligned}
& \text{Стандартная ошибка прогноза} = \\
& = S\sqrt{1+1/n} = \\
& = 2,94897\sqrt{1+1/8} = \\
& = 2,94897\sqrt{1,125} = \\
& = (2,94897)(1,06066) = \\
& = 3,12786
\end{aligned}$$

Для двустороннего 95% интервала предсказания t -значение из таблицы для $n=8$ будет равно $t = 2,365$. Интервал предсказания находится в пределах от

$$\bar{X} - t(\text{Стандартная ошибка прогноза}) = 8,125 - (2,365)(3,12786) = 0,728$$

до

$$\bar{X} + t(\text{Стандартная ошибка прогноза}) = 8,125 + (2,365)(3,12786) = 15,52.$$

Вы предполагаете, что сроки поставки имеют приблизительно нормальное распределение. Тогда 8 наблюдений представляют собой случайную выборку из теоретической совокупности "типичных сроков поставки" и время выполнения следующего заказа на поставку является случайно выбранным из этой же генеральной совокупности. Окончательная формулировка утверждения об интервале предсказания будет следующей.

"Мы на 95% уверены, что время выполнения следующего заказа на поставку будет лежать в пределах от 0,7 до 15,5 дней".

Почему этот интервал предсказания имеет такой широкий диапазон? Диапазон отражает неопределенность ситуации. В последних 8 поставках время выполнения сильно изменялось. Естественно, это затруднило составление точного прогноза.

Если вы хотите убедиться, что следующее время поставки будет не слишком большим, вам следует построить односторонний интервал предсказания, используя $t = 1,895$, взятое из колонки таблицы для одностороннего 95 процентного доверительного интервала. В этом случае верхняя граница будет равна

$$\bar{X} + t_{\text{односторонний}}(\text{Стандартная ошибка прогноза}) = 8,125 + (1,895)(3,12786) = 14,05229.$$

Затем можно сформулировать следующее утверждение об одностороннем интервале предсказания.

Мы на 95% уверены, что время выполнения следующего заказа на поставку будет не больше 14 дней.

Если вас устроит 90 процентный односторонний интервал предсказания, то верхняя граница (при $t = 1,415$) будет следующей:

$$\bar{X} + t_{\text{односторонний}}(\text{Стандартная ошибка прогноза}) = 8,125 + (1,415)(3,12786) = 12,55092.$$

Тогда можно сформулировать следующее утверждение об одностороннем доверительном интервале.

Мы на 95% уверены, что время выполнения следующего заказа на поставку будет не больше 12,6 дней.

9.6. Дополнительный материал

Резюме

Процесс обобщения данных выборки, который приводит к вероятностным утверждениям обо всей генеральной совокупности, называют **статистическим выводом**. **Доверительным интервалом** называют интервал, рассчитанный из данных таким образом, что существует *известная вероятность* включения интересующего вас (неизвестного) параметра генеральной совокупности в интервал, и эта вероятность интерпретируется с точки зрения случайного эксперимента, начинающегося с извлечения случайной выборки. Вероятность того, что параметр совокупности будет принадлежать доверительному интервалу называют **уровнем доверительности**, который обычно устанавливают равным 95%, хотя часто используют и другие уровни — 90, 99, 99,9%. Чем выше уровень доверительности, тем шире (а значит, и менее полезен) доверительный интервал. Приблизительная обобщенная формулировка утверждения о доверительном интервале имеет следующий вид.

Мы уверены на 95%, что значение параметра генеральной совокупности находится между значением оценки минус две стандартные ошибки оценки и значением оценки плюс две стандартные ошибки оценки.

Это утверждение основано на том факте, что при нормальном распределении с вероятностью 0,95 следует ожидать значения на расстоянии 1,960 (приблизительно 2) стандартного отклонения от среднего.

Формулировка утверждения о двустороннем 95% доверительном интервале для среднего генеральной совокупности имеет следующий вид.

Мы уверены на 95%, что среднее генеральной совокупности μ находится между $\bar{X} - t S_{\bar{X}}$ и $\bar{X} + t S_{\bar{X}}$, где значение t берется из t -таблицы.

Для биномиального распределения (при больших n) получаем такую формулировку утверждения о доверительном интервале.

Мы уверены на 95%, что доля интересующего нас свойства в генеральной совокупности π находится между $p - tS_p$ и $p + tS_p$, где значение t берется из t -таблицы.

Чтобы получить доверительный уровень, отличный от 95%, следует просто при построении доверительного интервала использовать соответствующее значение. **t-таблицу** используют для коррекции дополнительной неопределенности, обусловленной тем, что вместо неизвестного точного значения изменчивости генеральной совокупности используют оценку (стандартную ошибку). Когда вы работаете с однократной выборкой размера n , число **степеней свободы**, равное $n - 1$, представляет собой количество независимых элементов информации, использованных при вычислении стандартной ошибки (поскольку при вычислении стандартного отклонения из наблюдаемых значений вычитают среднее). Если известно точное значение стандартной ошибки, используют t -значение для бесконечного числа степеней свободы.

Для того чтобы **использование доверительного интервала было корректным**, необходимо выполнение двух следующих условий: (1) данные должны представлять собой случайную выборку из рассматриваемой генеральной совокупности,

(2) измеренные значения должны подчиняться нормальному распределению. Первое условие гарантирует, что данные правильно представляют неизвестный параметр, а второе дает основание использовать t -таблицу для вычисления вероятности. Поскольку в практической деятельности вычисление доверительного интервала основано главным образом на выборочном среднем \bar{X} , центральная предельная теорема позволяет смягчить условие 2, поэтому даже для умеренно асимметричного распределения при достаточно большом объеме выборки можно считать это условие выполненным.

Основанием для заявления об уверенности 95% или о доверительности 95% является то, что, как только значения для доверительного интервала вычислены, они уже не случайны; событие, имеющее вероятность 0.95, должно содержать случайность построения выборки. Интерпретация, основанная на относительной частоте, заключается в том, что если многократно повторять процесс взятия выборки и вычислять каждый раз доверительный интервал, то приблизительно 95% этих случайных известных интервалов будут включать конкретное неизвестное среднее значение генеральной совокупности. Подобным же образом из вычисленных вами в течение всей жизни при выполнении необходимых условий доверительных интервалов приблизительно 95% будут правильными (т.е. будут содержать неизвестный параметр) и приблизительно 5% будут ошибочными. Однако, вообще говоря, вы не будете знать, какие из них верны, а какие нет.

Односторонний доверительный интервал с известной доверительностью указывает, что среднее генеральной совокупности либо *не меньше*, либо *не больше* некоторого вычисленного значения. Вы вычисляете граничное значение для одностороннего доверительного интервала таким же образом, как и для двустороннего интервала, только заменяете t -значение для двустороннего интервала на t -значение для одностороннего интервала и выбираете граничную точку интервала так, чтобы построенный односторонний интервал включал выборочное среднее \bar{X} . При использовании одностороннего интервала вы должны быть уверены, что независимо от *поведения данных* вы будете использовать односторонний интервал с той же стороны (т.е. открытый в сторону больших значений или открытый в сторону меньших значений). В противном случае использование одностороннего доверительного интервала некорректно. При наличии сомнений лучше использовать двусторонний интервал. Утверждение об одностороннем доверительном интервале формулируется следующим образом.

Мы уверены на 95%, что среднее генеральной совокупности *не меньше*, чем $\bar{X} - t_{\text{односторонний}} S_{\bar{X}}$

или

Мы уверены на 95%, что среднее генеральной совокупности *не больше*, чем $\bar{X} + t_{\text{односторонний}} S_{\bar{X}}$.

Интервал предсказания позволяет использовать данные выборки для предсказания с известной вероятностью значения нового наблюдения при условии, что это новое наблюдение получено тем же способом, что и предшествующие. В качестве меры неопределенности здесь используется **стандартная ошибка предсказания** $S\sqrt{1+1/n}$, мера изменчивости расстояния между средним значением выборки и новым наблюдением. Интервал предсказания строят тем же способом, что и доверительный интервал; просто заменяют стандартную ошибку среднего

на стандартную ошибку предсказания. Формулировка утверждения об интервале предсказания (двустороннем) для значения нового наблюдения будет следующей.

Мы уверены на 95%, что новое наблюдение будет находиться между $\bar{X} - tS\sqrt{1+1/n}$ и $\bar{X} + tS\sqrt{1+1/n}$.

Формулировка утверждения об интервале предсказания (одностороннем) для значения нового наблюдения будет такой.

Мы уверены на 95%, что новое наблюдение будет не меньше, чем $\bar{X} - t_{\text{односторонний}}S\sqrt{1+1/n}$

или

Мы уверены на 95%, что новое наблюдение будет не больше, чем $\bar{X} + t_{\text{односторонний}}S\sqrt{1+1/n}$.

Выбирая соответствующие t -значения из таблицы, можно строить интервалы предсказания для уровней доверительности, отличных от 95%. Необходимо помнить, что доверительный интервал дает информацию о среднем генеральной совокупности, в то время как интервал предсказания дает информацию о единственном новом наблюдении, случайно выбранном из той же генеральной совокупности.

Основные термины

- Статистический вывод (statistical inference), 397
- Доверительный интервал (confidence interval), 397
- Уровень доверительности (confidence level), 397
- t -таблица (t-table), 401
- Степени свободы (degrees of freedom), 403
- Необходимые для построения доверительного интервала предварительные условия (assumptions required for the confident interval), 413
- Односторонний доверительный интервал (one-sided confidence interval), 419
- Интервал предсказания (prediction interval), 423
- Стандартная ошибка предсказания (standard error for prediction), 423

Контрольные вопросы

1. В чем статистический вывод выходит за пределы простого обобщения данных?
2. Какую дополнительную информацию о генеральной совокупности дает доверительный интервал по сравнению со значением оценки параметра?
3. Какое свойство нормального распределения приводит к тому, что в обобщенной формулировке утверждения о доверительном интервале фигурирует число 2 (или 1,96)?
4. Почему правильно говорить: “Мы на 95% уверены, что среднее совокупности находится между \$15,85 и \$19,36”, а не “Вероятность того, что среднее совокупности находится между \$15,85 и \$19,36, составляет 0,95”?
5. Почему для двустороннего 95%-ного доверительного интервала t -таблица содержит значения, большие чем 1,960?

6. а) Чему равно число степеней свободы для однократной выборки размером n ?
 б) Почему при вычислении числа степеней свободы вычитают единицу?
 в) Как определяется число степеней свободы, если известно точное значение стандартной ошибки?
7. а) Какие еще уровни доверительности, кроме 95%, используют достаточно часто?
 б) Чем отличается вычисление 99% доверительного интервала от 95% интервала?
 в) Какой из двух доверительных интервалов больше: двусторонний 99% или двусторонний 95% ?
8. а) Укажите два условия, которые должны выполняться для корректного построения доверительного интервала.
 б) Для каждого из условий приведите пример неверного результата в случае нарушения этого условия.
 в) Каким образом центральная предельная теорема помогает выполнению одного из этих условий?
 г) При каких обстоятельствах центральная предельная теорема не гарантирует выполнения второго условия?
9. а) Дайте основанную на относительной частоте интерпретацию корректности доверительного интервала.
 б) Дайте интерпретацию корректности большого количества доверительных интервалов с точки зрения “жизненных достижений”.
10. а) Почему односторонний доверительный интервал всегда должен включать среднее выборки?
 б) Должен ли односторонний доверительный интервал всегда включать в себя среднее генеральной совокупности?
11. а) Какой дополнительный критерий должен удовлетворяться для корректности одностороннего доверительного интервала (в дополнение к двум условиям для двустороннего доверительного интервала)?
 б) При наличии сомнений, какой доверительный интервал вы будете использовать: односторонний или двусторонний?
12. а) Какова разница между интервалом предсказания и доверительным интервалом?
 б) Какой тип интервала вы использовали бы, чтобы узнать о среднем значении привычных расходов вашего типичного покупателя?
 в) Какой тип интервала вы использовали бы, чтобы узнать о среднем значении привычных расходов отдельного покупателя?
13. а) Как вычисляют стандартную ошибку предсказания?
 б) Почему стандартная ошибка прогноза больше стандартного отклонения S ?

14. а) Что нужно изменить в вычислениях двустороннего 95% интервала предсказания, чтобы получить вместо него двусторонний 99% интервал предсказания?
- б) Что нужно изменить в вычислениях двустороннего 95% интервала предсказания, чтобы получить вместо него односторонний 95% интервал предсказания?
- в) Что нужно изменить в вычислениях двустороннего 95% интервала предсказания, чтобы получить вместо него односторонний 90% интервал предсказания?

Задачи

1. Ваша сельскохозяйственная фирма собирается приобрести некоторый большой участок пригодной для обработки земли. Для принятия решения необходимо изучить плодородие земли на этом участке. Случайная выборка из 62 небольших участков демонстрирует среднюю урожайность 103,6 бушелей кукурузы с акра со стандартным отклонением 9,4 бушеля с акра. Постройте двусторонний 95% доверительный интервал для среднего урожая, собранного со всего большого участка земли, возможность покупки которого изучается.
2. Ваша компания производит и распространяет замороженные пищевые продукты. Одна упаковка данного продукта должна иметь вес — 24,5 унции. Была взвешена случайная выборка из дневной продукции, и результаты оказались следующими: средний вес 24,41 унции, стандартное отклонение — 0,11 унции, размер выборки — 5 упаковок. Постройте двусторонний 95% доверительный интервал для среднего веса всех упаковок, выпущенных за этот день.
3. Ваша больница ведет переговоры с фирмой, занимающейся медицинским страхованием. Эта фирма намерена уменьшить суммы оплаты пребывания в больнице. Для конкретного вида лечения страховая фирма хотела бы уменьшить сумму оплаты пребывания в больнице на \$300, а также снизить время пребывания пациентов в больнице на один день. Чтобы определить влияние этих действий на больничные расходы, проанализировали случайную выборку из 50 пациентов, недавно прошедших такой курс лечения. В случае их выписки из больницы на день раньше средняя экономия составила бы \$322,44 со стандартным отклонением \$21,71. Постройте двусторонний 95% доверительный интервал для среднего размера экономии на одного пациента для большей совокупности недавно прошедших лечение пациентов.
4. Ваш отдел технического контроля качества проанализировал содержание 20 случайно отобранных бочек с материалами, которые используют при изготовлении пластикового садового инструмента. Получены следующие результаты: среднее — 41,93 галлона пригодного к употреблению материала в каждой бочке со стандартной ошибкой 0,040 галлона на бочку. Определите двусторонний 95% доверительный интервал для среднего генеральной совокупности.

5. Замерена интенсивность светового потока восьми карманных фонариков. Определите соответствующие t -значения из t -таблицы для вычисления следующих доверительных интервалов.
- а) Двусторонний, доверительность 95%.
 - б) Двусторонний, доверительность 99%.
 - в) Двусторонний, доверительность 99,9%.
 - г) Двусторонний, доверительность 90%.
6. Проведены исследования издержек для 21 производственной ситуации. Определите соответствующие t -значения из t -таблицы для вычисления следующих доверительных интервалов.
- а) Двусторонний, доверительность 95%.
 - б) Двусторонний, доверительность 99%.
 - в) Двусторонний, доверительность 99,9%.
 - г) Двусторонний, доверительность 90%.
7. Получены данные о реакции на вакцинацию у 1859 человек. Определите соответствующие t -значения из t -таблицы для вычисления следующих доверительных интервалов.
- а) Двусторонний, доверительность 95%.
 - б) Двусторонний, доверительность 99%.
 - в) Двусторонний, доверительность 99,9%.
 - г) Двусторонний, доверительность 90%.
8. Сорок восемь посетителей ресторана поставили оценки первому блюду. Определите соответствующие t -значения из t -таблицы для вычисления следующих доверительных интервалов.
- а) Двусторонний, доверительность 95%.
 - б) Двусторонний, доверительность 99%.
 - в) Двусторонний, доверительность 99,9%.
 - г) Двусторонний, доверительность 90%.
9. Собранные данные об урожайности имеют 17 степеней свободы. Определите соответствующие t -значения из t -таблицы для вычисления следующих доверительных интервалов.
- а) Двусторонний, доверительность 95%.
 - б) Двусторонний, доверительность 99%.
 - в) Двусторонний, доверительность 99,9%.
 - г) Двусторонний, доверительность 90%.
10. Исследованы предпочтения потребителей в ситуации, для которой известна стандартная ошибка. Определите соответствующие t -значения из t -таблицы для вычисления следующих доверительных интервалов.
- а) Двусторонний, доверительность 95%.

- б) Двусторонний, доверительность 99%.
 - в) Двусторонний, доверительность 99,9%.
 - г) Двусторонний, доверительность 90%.
 - д) Односторонний, доверительность 95%.
 - е) Односторонний, доверительность 99%.
11. Опрошена случайная выборка из восьми покупателей с целью определения количества персональных компьютеров, которые они планируют заказать в следующем году. Получены следующие результаты: 22, 18, 24, 47, 64, 32, 45 и 35. Вас интересует информация о большей генеральной совокупности, которую представляют опрошенные покупатели.
- а) Найдите обычную обобщенную меру изменчивости для отдельных покупателей.
 - б) Насколько приблизительно отличается значение среднего выборки от среднего генеральной совокупности?
 - в) Определите 95% доверительный интервал для среднего генеральной совокупности.
 - г) Определите 99% доверительный интервал для среднего генеральной совокупности.
12. Ваша пекарня выпекает батоны хлеба, вес которых, как указано на этикетке, составляет 1 фунт. Ниже приведены веса батонов, случайно отобранных из сегодняшней выпечки.
1,02; 0,97; 0,98; 1,10; 1,00; 1,02; 0,98; 1,03; 1,05; 1,02; 1,06.
Определите 95% доверительный интервал для среднего веса всех батонов из сегодняшней выпечки.
13. Маркетинговое исследование на основе выборочного опроса 483 человек показало, что в следующем году один человек будет тратить на ваше изделие в среднем \$15,48. Стандартное отклонение выборки составило \$2,52. Определите двусторонний 95% доверительный интервал для средних трат одного человека в будущем году для большей генеральной совокупности.
14. Последний опрос 252 потребителей, случайно отобранных из базы данных, содержащей 12 861 потребителей, показал, что 208 из них удовлетворены уровнем сервиса. Определите 99% доверительный интервал для процента удовлетворенных сервисом людей во всей базе данных.
15. Анализ выборки размером 258 человек, случайно отобранных из населения города численностью 750 339 человек, показал, что 165 из них поддерживают новый проект общественных работ. Определите 99,9% доверительный интервал для процента поддерживающих проект людей во всем городе.
16. Из 763 случайно отобранных человек 152 не знакомы с вашей продукцией.
- а) Оцените процент людей в генеральной совокупности (из которой была извлечена эта выборка), не знакомых с вашей продукцией.
 - б) Определите стандартную ошибку оценки, вычисленной в п. "а".

- в) Определите двусторонний 95% доверительный интервал для процента таких людей в генеральной совокупности.
- г) Определите односторонний 99% доверительный интервал, согласно которому процент таких людей в совокупности не превышает определенное значение.
- д) Почему этот статистический вывод является приблизительно корректным, даже если распределение в генеральной совокупности не является нормальным?
17. При проведении общенационального опроса требуется, чтобы границы ошибки не превышали 3 процентные единицы в каждом направлении (т.е. плюс или минус) при уровне доверительности 95%.
- а) Проверьте выполнение этого требования для конкретного случая, вычислив произведение t -значения из таблицы на стандартную ошибку биномиально распределенной доли p для ситуации, когда 309 из 1105 зарегистрированных избирателей утверждают, что они поддержат на выборах определенного кандидата.
- б) Постройте 95% доверительный интервал для выраженной в процентах доли зарегистрированных избирателей, которые поддержат определенного кандидата, как это указано в п. "а".
18. При проведении общенационального опроса граница ошибки не должна превышать 4,3 процентные единицы для вопросов, заданных половине выборки.
- а) Проверьте выполнение этого требования для конкретного случая, вычислив произведение t -значения из таблицы на стандартную ошибку биномиально распределенной доли p для случая, когда 46% из 553 зарегистрированных избирателей-женщин утверждают, что они поддержат на выборах определенного кандидата.
- б) Постройте 95% доверительный интервал для выраженной в процентах доли зарегистрированных избирателей-женщин, которые поддержат определенного кандидата, как это указано в п. "а".
19. В результате опроса 15 инженеров, работающих в отраслях металлообрабатывающей промышленности, которые пребывают в своей должности от одного до трех лет и являются членами Американского общества контроля качества продукции, было установлено, что их средняя зарплата составляет \$37 496 в год.¹⁴ Будем считать, что это случайная выборка со стандартным отклонением \$9000.
- а) Определите 95% доверительный интервал для средней зарплаты в генеральной совокупности.
- б) Определите 99% доверительный интервал для средней зарплаты в генеральной совокупности.
- в) Закончите следующее предложение: "Мы на 95% уверены, что средняя зарплата в генеральной совокупности по крайней мере составляет _____".

¹⁴ Quality Progress, November 1995, p. 58.

20. В результате опроса девяти менеджеров транспортной и аэрокосмической отрасли, которые работают в своей должности более 20 лет и являются членами Американского общества контроля качества продукции, было установлено, что их средняя зарплата составляет \$67 333 в год.¹⁵ Будем считать, что это случайная выборка со стандартным отклонением \$19 000.

а) Определите 95% доверительный интервал для средней зарплат в генеральной совокупности.

б) Определите 99% доверительный интервал для средней зарплат в генеральной совокупности.

в) Закончите следующее предложение: “Мы на 95% уверены, что средняя зарплата в генеральной совокупности по крайней мере составляет _____”.

21. Ваша фирма хочет нанять опытного квалифицированного менеджера по контролю качества продукции. Считая, что такой менеджер отобран случайно из генеральной совокупности, представленной в предыдущей задаче, закончите следующее предложение: “Мы на 99% уверены, что зарплата выбранного менеджера должна составлять по крайней мере _____”.

22. Многие потребители считают, что они смогут сэкономить деньги, делая покупки в новом магазине под названием SuperMall. В табл. 9.6.1 показаны результаты исследования случайно отобранных товаров. Определите 95% доверительный интервал для среднего значения (в долларах) сэкономленных средств в генеральной совокупности, которую представляют эти товары.

23. Исходя из приведенных ниже дневных процентных колебаний индекса S&P 500 фондового рынка в июле 1995 года определите доверительный интервал для среднего дневного изменения в генеральной совокупности. (Это, строго говоря, не является случайной выборкой из генеральной совокупности. Однако теория случайных изменений фондового рынка предполагает, что колебания рынка должны вести себя как случайная выборка. Генеральная совокупность будет представлять собой все дневные изменения на рынке, которые могли бы произойти в условиях, преобладающих в данное время.)

0,43%; 0,03%; 1,23%; 0,43%; 0,15%; -0,43%; 1,10%; 0,02%;
-0,20%; 0,51%; -0,76%; -1,34%; 0,46%; 0,01%; 0,54%; 0,80%;
0,09%; 0,64%; -0,41%; -0,15%

24. Во время однонедельного эксперимента по всей стране в магазинах вашей фирмы был изменен способ демонстрации рекламы товара внутри магазина. В результате объем продаж для этих товаров увеличился (по сравнению с предыдущей неделей) в среднем на \$441,84 со стандартным отклонением \$68,91. В эксперименте участвовало 18 магазинов. (Аналогичные ситуации были исследованы компанией Bennett-Chaikin, Inc., о чем сообщалось в рекламе корпорации Menasha Corporation в *Advertising Age*, August 21, 1995, p. 17.)

¹⁵ Quality Progress, November 1995, p. 61.

Таблица 9.6.1. Цены в магазине SuperMall и в других местах на различные товары


Товар	Цена в SuperMall, дол.	Цена в других местах, дол.	Экономия при покупке в SuperMall, дол.
Levi's	27,99	37,99	10,00
Reebok Walk Leader	64,99	69,99	5,00
Farberware Omelette Pan	89,99	89,99	0,00
Chocolate Meal Replacement Shake	12,99	12,99	0,00
Kolcraft Baby Stroller	109,75	119,99	10,24
Wilson's Lambskin Jacket	199,00	199,00	0,00
Joop!	47,95	55,00	7,05
Eureka Tent	239,99	239,99	0,00
Geoffey Beene Shirt	24,99	39,50	14,51
JVC VCR	398,87	386,97	-11,90
Nike Basketball Shoes	79,99	79,99	0,00
Spalding NBA Ball	69,99	69,99	0,00
Smoothe as Silk	12,99	12,99	0,00
SBA Coffee	9,50	8,95	-0,55
Barbie House	64,99	64,99	0,00
"Shameless" CD	15,99	15,99	0,00

Данные взяты из *The Seattle Times*, September 22, 1995, p. A16.

а) Определите 95% доверительный интервал для среднего значения увеличения продаж в генеральной совокупности.

б) Закончите следующее предложение: "Мы на 95% уверены, что среднее значение объема продаж в генеральной совокупности увеличится не меньше чем на _____".

в) Менеджер одного из ваших магазинов хотел бы оценить возможное увеличение продаж. Этот магазин не вошел в исследование. Предполагая, что условия в этом магазине аналогичны условиям в тех магазинах, которые принимали участие в эксперименте, закончите следующее предложение: "Мы на 95% уверены, что при изменении способа демонстрации рекламы увеличение недельного объема продаж в данном магазине составит от ___ до ___".

25. В табл. 9.6.2 представлены значения доходности акций, которые предлагают брокерские фирмы. 

а) Вычислите среднее и дайте краткую интерпретацию найденному значению.

б) Вычислите стандартное отклонение и дайте краткую интерпретацию найденному значению.

в) Вычислите стандартную ошибку и дайте краткую интерпретацию найденному значению.

Таблица 9.6.2. Доходности акций, предложенных брокерскими фирмами

Фирма	Доходность, %	Фирма	Доходность, %
Thomson McKinnon	15,8	A. G. Edwards	1,2
Shearson Lehman	5,5	PaineWebber	0,4
Prudential-Bache	4,7	E. F. Hutton	0,3
Merrill Lynch	2,9	Dean Witter	-2,8
Drexel Burnham	2,5	Kidder Peabody	-7,1

Взято из J. R. Dorfman "Do Brokers' Stock Picks Perform Well?", *The Wall Street Journal*, February 5, 1988, p. 17.

- г) Определите двусторонний 95% доверительный интервал среднего значения доходности акций, предложенных аналогичными брокерскими фирмами в течение этого же периода времени, рассматривая имеющийся набор данных как случайную выборку из этой идеализированной совокупности.
- д) Определите двусторонний 90% доверительный интервал и сравните его с 95% доверительным интервалом.
- е) Найдите односторонний 99% доверительный интервал для случая, когда среднее значение доходности *не меньше* чем некоторое значение.
- ж) Предположим, анализируя данные, вы видите, что идет уменьшение среднего значения доходности, и хотите сформулировать утверждение об одностороннем доверительном интервале о том, что среднее значение доходности *не больше* чем некоторая величина (вместо вашего ответа в п. "е"). В этом случае и при использовании одной и той же таблицы данных будет ли сформулированное вами в п. е утверждение о доверительном интервале корректным? Почему да или почему нет? Если нет, то что нужно делать вместо этого?
26. Предвыборный опрос 921 случайно отобранного избирателя показал, что ваш кандидат впереди с 52,4% голосов.
- а) Определите двусторонний 95% доверительный интервал для процента голосов за вашего кандидата в генеральной совокупности. **A**
- б) Поскольку вы давно симпатизируете этому кандидату и желаете ей победы на выборах, вас интересует только информация о том, будет ли за нее отдано не менее некоторого процента голосов избирателей. Будет ли правильным в данном случае строить односторонний доверительный интервал?
- в) С учетом информации из п. "б" постройте подходящий односторонний 95% доверительный интервал.
- г) Постройте подходящий односторонний 90% доверительный интервал.
- д) Опишите в одном абзаце текста, как эти доверительные интервалы проливают свет на шансы вашего кандидата. В частности, насколько больше вы знаете теперь в сравнении с одним числом 52,4%?
27. Исследование рынка, основанное на выборочном опросе 400 человек, показало, что люди готовы тратить в следующем году на покупку вашей про-

дукции в среднем \$2,34 (на одного человека). Стандартное отклонение выборки составило \$0,72. Определите двусторонний 95% доверительный интервал для средних трат одного человека в будущем году для большей генеральной совокупности.

28. К вашему удивлению, опрос ваших клиентов показал, что из случайно отобранных 200 клиентов 42 недовольны послепродажной поддержкой и обслуживанием.

а) Вычислите основные статистики: размер выборки, n ; выборочный процент, p ; стандартную ошибку, S_p .

б) Постройте двусторонний 95% доверительный интервал для процента недовольных среди *всех* ваших клиентов (т.е. не только тех, кто был опрошен).

в) Ваша генеральная совокупность состоит из 28 209 клиентов. Преобразуйте проценты, представляющие граничные точки доверительного интервала в п. “б”, в количество людей в генеральной совокупности. Сформулируйте и объясните ваш результат как доверительный интервал для *количества* недовольных потребителей в генеральной совокупности.

29. Выборка 93 рулонов листовой стали показала, что средняя длина рулона составляет 101,37 метра со стандартным отклонением 2,67 метра.

а) Дайте словесную интерпретацию стандартного отклонения; в частности, каким образом стандартное отклонение измеряет изменчивость?

б) Найдите стандартную ошибку. Дайте словесную интерпретацию найденного значения и сравните ее с представленной в п. “а” интерпретацией стандартного отклонения.

в) Определите двусторонний 95% доверительный интервал для средней длины рулона стали в большей генеральной совокупности. Кратко опишите полученное значение.

г) Определите двусторонний 95% интервал предсказания для длины рулона, который будет произведен следующим. Кратко опишите полученный результат и сравните этот интервал предсказания с доверительным интервалом из п. “в”. В частности, почему интервал предсказания настолько шире доверительного интервала?

д) Ваша совесть требует, чтобы рулоны стали были гарантированно *не меньше* определенной длины. Заставит ли эта информация вас вычислять односторонний интервал? Почему да или почему нет?

е) Определите соответствующий односторонний 99% доверительный интервал и объясните полученный результат.

ж) Определите соответствующий односторонний 99% интервал предсказания и объясните полученный результат.

30. В процессе написания брошюры, описывающей быстродействие новой компьютерной системы, вы измеряете, сколько времени необходимо специальной тестовой программе для обработки некоторой базы данных. Ввиду того что по мере добавления, удаления или изменения записей физическое раз-

мещение базы данных на диске постоянно меняется, в результатах выполнения измерительного теста наблюдается некоторая вариация. Ниже приведены значения времени (в минутах) для 14 независимых повторений теста.

5, 6, 8, 11, 5, 8, 11, 10, 6, 10, 5, 9, 5, 5

а) Определите двусторонний 95% доверительный интервал и опишите полученное значение с точки зрения среднего значения производительности системы при ее массовом выпуске.

б) Определите подходящий односторонний 95% доверительный интервал исходя из того, что вы хотите показать, насколько быстро работает система (лучшими являются более низкие значения).

в) Определите соответствующий односторонний 90% доверительный интервал.

г) Определите соответствующий односторонний 99% доверительный интервал.

д) Напишите краткий отрывок для рекламной брошюры, описывающий один (или несколько) из полученных результатов. Будьте честны, но “покажите товар лицом” и пишите обычным доступным языком. Включите в качестве сносок дополнительную техническую информацию, чтобы более подготовленные читатели могли понять подробности сделанной вами работы.

31. Проведен анализ выборки горной породы, взятой из различных мест в некоторой шахте. Для каждой выборки вычислено значение “нормы прибыли”, представляющей собой прибыль (полученную после продажи выплавленного металла по текущей рыночной цене) как процент от стоимости добычи. Эта величина отражает затраты на добычу руды, ее переработку и доходность конечного продукта, все в конкретных экономических условиях. Можно рассматривать эти данные как случайную выборку, которая соответствует реальным условиям функционирования экономически жизнеспособного производства. Для экономической устойчивости необходимо, чтобы норма прибыли была достаточно высокой и оправдывала затраты на производство. Ниже приведены значения нормы прибыли, полученные для 13 выборок.

8,1%; 6,2%; 19,8%; -4,3%; 5,1%; 0,2%; -10,4%; 11,8%; 2,0%; 4,7%;
-3,2%; 8,9%; -6,2%

а) Вычислите основные статистики: n , среднее, стандартное отклонение и стандартную ошибку. Кратко опишите ситуацию, представив, что вы должны дать объяснение совету директоров.

б) Определите генеральную совокупность и среднее генеральной совокупности. Почему значение среднего генеральной совокупности является важным как для управления, так и для владельцев данной шахты?

в) Определите соответствующий односторонний 99% доверительный интервал. Письменно поясните полученное значение.


г) Подготовьте краткую докладную записку для высшего руководства, в которой опишите в общих чертах ситуацию и дайте рекомендации возможных действий.

32. Вас беспокоит брак при печати газеты. Никакие количественные измерения ранее не проводились, несмотря на то, что частые ошибки вынуждают выбрасывать много фунтов только что напечатанной продукции. Чтобы оценить серьезность проблемы и помочь в ее решении, если это будет оправдано, вы начали собирать данные. Вы будете предпринимать какие-то меры только в том случае, если количество брака достаточно велико. Взвешены бракованные экземпляры газеты для 27 отобранных утренних выпусков. Средний вес брака составил 273,1 фунта в день со стандартным отклонением 64,2 фунта.
- а) Подходит ли для данной ситуации построение одностороннего доверительного интервала? Почему да или почему нет?
 - б) Постройте наиболее полезный в данной ситуации односторонний 99% доверительный интервал. Почему вы выбрали именно такой односторонний интервал?
 - в) Выразите построенный доверительный интервал в фунтах за год, приняв, что в году 365 рабочих дней.
 - г) Определите односторонний 99% интервал предсказания для завтрашнего количества брака. Сравните полученный результат с результатом из предыдущего пункта.
33. На новой работе вы заключили девять контрактов на продажу продукции со средней ценой \$3782 и стандартным отклонением \$1290.
- а) Определите подходящую идеализированную совокупность, которую представляет данная выборка.
 - б) Если распределение цен продаж достаточно сильно асимметрично, можно ли строить обычный двусторонний 95% доверительный интервал? Почему да или почему нет?
 - в) Теперь предположим, что распределение цен только слегка асимметрично и не сильно отличается от нормального. Вычислите обычный двусторонний 95% доверительный интервал и поясните его с точки зрения долгосрочной перспективы вашей работы на этом месте. Отрадите как полезную информацию, представленную доверительным интервалом, так и ограничения доверительного интервала.
 - г) Найдите двусторонний 90% интервал предсказания для цены продаж следующего заключенного вами контракта, приняв, что условия останутся в основном теми же.
34. Анализ случайной выборки записей о 50 пациентах, недавно посетивших клинику, демонстрирует, что средняя стоимость одного визита к врачу составляет \$53,01 со стандартным отклонением \$16,48.
- а) Найдите 95% доверительный интервал для среднего и дайте его интерпретацию.
 - б) Найдите 99% доверительный интервал.
 - в) Найдите односторонний 95% доверительный интервал, определяющий по крайней мере некоторый уровень стоимости визита к врачу.

35. Найдите 95% доверительный интервал для суммы денег, которую ваши постоянные покупатели потратили на ваши товары в прошлом месяце, используя данные из задачи 2 главы 4 в качестве случайной выборки заказов покупателей.
36. Найдите 99% доверительный интервал для прочности хлопковой нити, используемой на ткацкой фабрике, используя данные задачи 15 главы 4.
37. Определите 99,9% доверительный интервал для веса шоколадного батончика до вмешательства в процесс производства, используя данные задачи 11 главы 5.
38. Используя данные задачи 11 главы 5, определите односторонний 95% доверительный интервал для веса шоколадного батончика после вмешательства в процесс производства, указав, что средний вес в генеральной совокупности *не превышает* определенное значение.
39. Из списка 729 человек, участвовавших в круизе, для опроса случайно отобрали 130. Из них 112 человек ответили, что они очень довольны обслуживанием во время круиза. Найдите 95% доверительный интервал для процента тех людей в генеральной совокупности, которые довольны предоставленным обслуживанием.
40. Используя данные о качестве сельскохозяйственной продукции из задачи 29 главы 8.
- а) Найдите 95% доверительный интервал для среднего качества продуктов во всей генеральной совокупности.
- б) Найдите 95% интервал предсказания для результатов измерения качества продукции в следующем наблюдении.
- в) Найдите 99% доверительный интервал для среднего качества продукции во всей генеральной совокупности.
- г) Найдите 99% интервал предсказания для результатов измерения качества продукции в следующем наблюдении.
41. Ниже приведено содержание кофеина (в миллиграммах) в случайно отобранных чашках кофе.
112,8; 86,4; 45,9; 110,3; 100,3; 93,3; 101,9; 115,7; 92,5; 117,3; 105,6; 81,6.
- а) Найдите односторонний 99% доверительный интервал для среднего значения генеральной совокупности содержания кофеина в чашке кофе, который утверждает, что “не менее чем ...”.
- б) Найдите односторонний 99% интервал предсказания для содержания кофеина в следующей чашке кофе, также утверждающего “не менее чем ...”.

Упражнения с использованием базы данных

Обратитесь к базе данных служащих, приведенной в приложении А.

1. Будем считать эту базу данных генеральной совокупностью. Рассмотрим следующую выборку пяти номеров служащих в этой базе данных: 24, 54, 17, 34 и 53. 

- а) На основе данных этой выборки найдите среднее, стандартное отклонение и стандартную ошибку для годовой заработной платы.
- б) Найдите 95% доверительный интервал для средней заработной платы во всей генеральной совокупности.
- в) Начертите график (аналогичный рис. 9.1.5), содержащий среднее выборки и доверительный интервал.
2. А теперь рассмотрим всю генеральную совокупность размеров заработной платы, что в реальной жизни сделать невозможно.
- а) Найдите среднее и стандартное отклонение генеральной совокупности и сравните их с выборочными оценками.
- б) Начертите для данной ситуации график (аналогичный рис. 9.1.1). Используйте $\sigma_{\bar{x}}$ как стандартное отклонение выборочного распределения.
- в) Попадает ли в этом случае среднее генеральной совокупности в доверительный интервал (из задачи 1)? Будет ли это среднее *всегда* находиться в этом интервале для всех случайных выборок? Почему да или почему нет?
3. Повторите упражнение 1, пп. “б” и “в”, для 99% доверительного интервала. Попадает ли среднее генеральной совокупности годовой зарплаты в этот интервал?
4. Повторите упражнение 1, пп. “б” и “в”, для 90% доверительного интервала. Попадает ли среднее генеральной совокупности годовой зарплаты в этот интервал?
5. Повторите упражнение 1, используя 95% доверительный интервал для следующей случайной выборки: номера служащих 4, 47, 45, 12 и 69. Дайте ответы на следующие вопросы:
- а) В реальной жизни что бы вы могли предпринять (если это возможно) в отношении того факта, что среднее генеральной совокупности не принадлежит доверительному интервалу?
- б) Постройте 99% и 99,9% доверительные интервалы. При каком уровне доверительности (если таковой существует) доверительный интервал будет достаточно большим, чтобы включать среднее генеральной совокупности?
6. Рассмотрим следующую случайную выборку из 15 номеров служащих в этой базе данных: 66, 37, 56, 11, 32, 23, 53, 43, 55, 25,7, 26, 36, 22 и 20.
- а) Найдите процент женщин в этой выборке.
- б) Найдите стандартную ошибку для процента женщин и поясните полученное значение.
- в) Какие у вас могли бы быть сомнения в использовании этой выборки и методов, изложенных в этой главе, при вычислении доверительного интервала для процента женщин?
7. Считая базу данных из приложения А случайной выборкой из более крупной генеральной совокупности, рассмотрим размер годовой зарплаты служащего.
- а) Найдите 95% доверительный интервал.
- б) Найдите 99% доверительный интервал.

8. Считая базу данных из приложения А случайной выборкой из более крупной генеральной совокупности, рассмотрим возраст служащего.
 - а) Найдите 95% доверительный интервал.
 - б) Найдите 90% доверительный интервал.
9. Считая базу данных из приложения А случайной выборкой из более крупной генеральной совокупности, рассмотрим стаж работы служащего.
 - а) Найдите 95% доверительный интервал.
 - б) Найдите 99,9% доверительный интервал.
10. Считая базу данных из приложения А случайной выборкой из более крупной генеральной совокупности, рассмотрим процент женщин. Постройте 95% доверительный интервал.
11. Считая базу данных из приложения А случайной выборкой из более крупной генеральной совокупности, рассмотрим процент служащих высокой квалификации (имеют уровень подготовки В или С). Постройте 99% доверительный интервал.
12. Рассматривая базу данных из приложения А как случайную выборку из более крупной генеральной совокупности
 - а) Найдите односторонний 95% доверительный интервал для средней годовой зарплаты в генеральной совокупности, который утверждает, что зарплата по крайней мере составляет некоторую сумму.
 - б) Найдите 99% односторонний доверительный интервал для п. “а”.
 - в) Найдите 95% односторонний доверительный интервал для среднего значения стажа работы служащего в генеральной совокупности, который утверждает, что стаж составляет *по крайней мере* определенное количество лет.
 - г) Найдите 99% односторонний доверительный интервал для п. “в”.
13. Рассматривая базу данных из приложения А как случайную выборку из идеализированной совокупности потенциальных сотрудников, которые могут быть приняты на работу
 - а) Определите 95% интервал предсказания для стажа работы сотрудника, который будет нанят следующим. Почему этот интервал намного шире, чем доверительный интервал для среднего значения стажа во всей генеральной совокупности?
 - б) Определите 95% интервал предсказания для возраста сотрудника, который будет нанят следующим.

Проекты

1. Получите оценку и стандартную ошибку этой оценки для *двух* показателей, важных для ваших интересов в бизнесе (используйте реальные данные или исходите из некоторых предположений). Для каждого из этих двух показателей постройте доверительный интервал и напишите небольшое пояснение к нему. В случае использования уровня доверительности, отличного от 95%, или в случае использования одностороннего доверительного интервала поясните, почему вы поступили именно так.

2. Получите оценку и стандартное отклонение этой оценки для *двух* показателей, важных для ваших интересов в бизнесе (используйте реальные данные или исходите из некоторых предположений). Для каждого из этих двух показателей постройте интервал предсказания и напишите небольшое пояснение к нему. В случае использования уровня предсказания, отличного от 95%, или в случае использования одностороннего интервала предсказания поясните, почему вы поступили именно так.
3. Найдите в Internet или в газете статью с результатами опроса общественного мнения. Напишите один абзац текста, посвященный одному из результатов опроса. Обязательно укажите объем выборки, процент и стандартную ошибку. Постройте сами двусторонний 95% доверительный интервал. Сравните ваши результаты с предельным значением ошибки опроса, если это значение указано в использованной вами статье.



Ситуация для анализа

Многообещающие результаты опроса относительно заказов фирменных товаров по каталогу

Недавно были подведены предварительные результаты исследования, касающегося проекта продаж фирменных товаров по каталогу, которые выглядят очень многообещающе! Средний планируемый размер заказа составил \$53,94, т.е. на \$15 больше ожидаемого. Руководитель группы от радости, вероятно, должен быть на седьмом небе: так как \$53,94 за каждый заказ, полученный из 1 300 000 потенциально возможных адресов, приводит к продажам со средней суммарной цифрой более 70 миллионов долларов!

При подготовке совещания в ваши обязанности входит оценить, насколько корректно было проведено данное исследование. Исходная докладная записка, кроме суммы в \$53,94, содержала мало деталей. Сделав несколько телефонных звонков, вы узнали, кто из служащих проделал основной объем работы в этом исследовании. Ниже приводится полученная вами информация. Случайная выборка была извлечена из соответствующей базы данных, содержащей адреса 600 000 обеспеченных людей, покупающих предметы роскоши по почте. По почте были разосланы 600 каталогов вместе с вопросниками. Также вы узнали, что 74 из этих 600 вопросников вернулись. В 9 из них было отмечено: “Да, я согласен до конца года заказать данные изделия на общую сумму _____ долларов”. Указаны были следующие суммы: \$7,97; \$12,05; \$29,27; \$228,26; \$2,28; \$7,25; \$114,39; \$31,64 и \$52,39.

Теперь вам известно, что в размерах заказа существует существенный разброс. 95% доверительный интервал для среднего начинается с \$3,10 и заканчивается \$104,79. Умножив эти числа на количество возможных покупателей (1 300 000), получаем в качестве границ суммы от 4 030 000 до 136 227 000 долларов. Поэтому, даже приняв во внимание случайность, можно реально надеяться получить такие объемы продаж. А может быть, нет?

Вопросы для обсуждения

1. Можно ли было умножить среднее значение суммы заказа (\$53,94) на количество (1 300 000) возможных покупателей?
2. Может, лучше умножить (как и предложено) граничные значения доверительного интервала на количество возможных покупателей?
3. А может, лучше умножать на количество адресов в основе выборки, которая использовалась для извлечения случайной выборки?
4. Может быть, вас еще что-либо смущает в этой ситуации?
5. Какой будет ваша наилучшая оценка с доверительными границами для потенциальных продаж по каталогу?